2020-04

A computational study on the role of solvents and conformational fluctuation of macromolecules towards drug design

Shadrack, Daniel Madulu

NM-AIST

A COMPUTATIONAL STUDY ON THE ROLE OF SOLVENTS AND CONFORMATIONAL FLUCTUATION OF MACROMOLECULES TOWARDS DRUG DESIGN

Daniel Madulu Shadrack

A Thesis Submitted in Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Life Sciences of the Nelson Mandela African Institution of Science and Technology

Arusha, Tanzania

April, 2020

ABSTRACT

Heat shock protein 90 (Hsp90) represents an important chemotherapeutic target in the treatment of various ailments including cancer and neurodegenerative diseases. The protein is responsible for controlling and regulating the growth of nearly 200 client proteins known to overexpress in tumour or cancer cells. Targeting Hsp90 and inhibiting its chaperone machinery function results in proteasome degradation of the client protein and hence treatment of the disease. In this thesis, different computational docking protocols, the role of water and conformational fluctuations in drug design for the discovery and identification of new Hsp90 inhibitors are reported. In particular, the sensitivity of different docking protocols to crystal structure with and without water, relaxed complex scheme (RCS) or ensemble-based structures holo and apo structures with and without water, the effects of including a different amount of water in the protein active site on the thermodynamics of ligand binding to protein structures are reported. There is sensitivity of results to different docking protocols, RCS lowers the binding energy in comparison to crystal structure, holo ensemble with strong ligand bound improves the docking results. Since biological activities of small molecules highly depends on the conformation, molecular structure, charge distribution and non-trivial response to solvents. The thesis further explored the role of different solvents *viz* polar protic, polar aprotic, and non-polar on the conformation of curcumin as a model drug/natural product. Well-tempered metadynamics (WT-MetaD), an enhanced sampling method employing OPLS-AA force field in an isobaric-isothermal (NPT) ensemble was used to investigate the related solvent effects. The orientation and conformational of curcumin was solvent dependent, the free energy for curcumin in solvents and vacuum portrayed a different behaviour. Curcumin exists in different configuration and conformations in different solvents. The *trans*-conformation was more stable in polar aprotic solvents capable of solubilizing curcumin whereas the *cis*-conformation was more stable in polar protic solvents i.e water where it has marginal solubility. Finally, the thesis reports on the influence of solvents on kinetics and residence time of drug unbinding in host-guest complexes. The effect of polar aprotic and polar protic solvents on kinetics and residence time of drug unbinding from a nanoparticle was investigated using chitosan-toussantine-A as a model system. WT-MetaD was used to study the kinetics and residence time. Results show that the kinetics and residence time of drug unbinding was affected by solvents. Slow unbinding kinetics of $k_{\text{off}} = 0.045\ \mu s^{-1}$ was observed for the system formulated with water, a polar protic solvent, while fast unbinding kinetics with $k_{\text{off}} = 1000\ \mu s^{-1}$ was observed in system formulated with DMSO solvent. Furthermore, the interaction of chitosan-toussantine-A complex in water was observed to be stable than in DMSO. The approaches used in this thesis pave the ways and can further be extended to investigate more problems in drug design ranging from protein-ligand interaction, solution conformation of small molecules and host-guest kinetics. Since the new reported small molecules as Hsp90 inhibitors are approved for other indication, the inhibitors are recommended for further pre-clinical and clinical testing as new Hsp90 inhibitors for cancer treatment.

DECLARATION

I, **Daniel Madulu Shadrack**, do hereby declare to the Senate of The Nelson Mandela African Institution of Science and Technology that this thesis is my own original work and that it has neither been submitted nor concurrently submitted for a degree or similar award in any other institution.

Name and Signature of Candidate	Date

This declaration is confirmed by:

Prof. Ali A. Hassanali	Date
Prof. Hulda S. Swai	Date

COPYRIGHT

CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance by The Nelson Mandela African Institution of Science and Technology a thesis titled: A computational study on the role of solvents and conformational fluctuation of macromolecules towards drug design, submitted by Daniel M Shadrack (P189/T.16), in fulfillment of the requirements for the degree of PhD in Life Sciences of The Nelson Mandela African Institution of Science and Technology.

Approved by:

Prof. Ali A. Hassanali	Date
Prof. Hulda S. Swai	Date

ACKNOWLEDGEMENT

First of all, I am thankful to the almighty God, for granting wisdom, good health and life.

The success and completion of this thesis would be impossible without the help and guidance of many people who have made to appear as it is now. It is, therefore, a great pleasure to thank many people that supported me during my PhD study. I would like to express my sincere gratitude to the Nelson Mandela African Institution of Science and Technology (NM-AIST), Tanzania, the *Abdus Salam* International Centre for Theoretical Physics (ICTP), Trieste, Italy and to my dedicated supervisors: Prof. Ali A. Hassanali (ICTP, Trieste, Italy) and Prof. Hulda S. Swai (NM-AIST, Arusha, Tanzania) for their support, guidance and constant encouragement throughout the journey of my PhD study. In particular, I thank you Ali Hassanali for constant guidance and tireless support you provided and introducing me not only to a dynamic and learning /research environment but also into research area I dreamed to do ...., indeed you are an instrument in my carrier lifetime, God bless you. Hulda Swai, I thank you for always availing the support I needed in my PhD journey, it was not easy but through the support you give to me, I have made it. I further express my sincere gratitude to Prof. Alessandro Laio (SISSA), for his constant and generous support, time, advice and readiness to provide any help during my study period.

I further extend my appreciations to CREATES, Director's office of the *Abdus Salam* International Centre for Theoretical Physics, and in particular Prof. Fernando Quevedo for financial support during my entire study period and for availing all computational support and facilities during my PhD study at ICTP. I further, thank my employer, the St. John's University of Tanzania (SJUT) for granting a study leave and support during my PhD study period.

My special appreciation also goes to all current and previous postdoc and PhD students under the group of Prof. Ali A. Hassanali and the Condensed Matter and Statistical Physics section (CMSP), ICTP, Trieste, Italy for the fruitful discussions we had during my entire period at ICTP. In particular, I thank Dr. Narjes Ansari., Dr. Emiliano Polli, Dr. Kwang Hyok Jong, Dr. Mohammed Nawaz and everyone I interacted with. I also express my thanks to my collaborators Dr. Lucy Kiruri (Kenyatta University, Nairobi, Kenya) and Dr. Stephen Samwel Nyandoro (University of Dar es Salaam, Tanzania). I also sincerely thank Profs., Tatiana Pogrebnaya and Alexander Pogrebnoi (NM-AIST, Arusha, Tanzania) for constant encouragement during my PhD study period. I further thank all my colleagues and friends at the NM-AIST for their support, I cannot mention you all, at least I mention a few, Geradius Deogratias Kikumi, all the students under Profs. Tatiana Pogrebnaya and Alexanders Pogrebnoi (NM-AIST), CREATES students and all COCSE students who we interacted in the COCSE Lab. Our meetings and interactions are so fruitful, thank you all. I also thank all the technical person in the HPC section,

Mr. Adam Mawenya, Ms. Joyce Martin and Ms. Leah Gonda for providing all the supports I needed in my study.

Last but not least, my special thanks also goes to my family members: my parents Mr. and Mrs. Peter M. Makungu, brothers, sisters, sibling and my lovely one, for their encouragement, prayers, love and all kind of support in my life.

DEDICATION

I dedicate this thesis to all my family members and my PhD supervisors

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
CVs	Collective variables
ff	Force fields
Hsp90	Heat shock protein 90
HSPs	Heat shock proteins
HSF-1	Heat shock factor1
HSR	Heat shock response
k_B	Boltzmann constant
MetaD	Metadynamics
MD	Molecular dynamics
MM-PBSA	Molecular mechanics Poisson-Boltzmann surface area
NVT	Constant number of particles, volume and temperature
NPT	Constant number of particles, pressure and temperature
OPLS-AA	Optimized potential for liquid systems-all atom
OPLS-UA	Optimized potential for liquid systems-united atom
PBC	Periodic boundary conditions
PME	Particle Mesh Ewald
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
TI	Thermodynamics integration
TouA	Toussantine-A
WT-MetaD	Well-tempered metadynamics

CHAPTER ONE

INTRODUCTION

1.1 Background of the problem

1.1.1 Brief introduction of heat shock proteins (HSPs) and their role in cancer treatment

It is well acknowledged that when cells are exposed to environmental or physiological stress, such as ultraviolet light, mechanical stress, anoxia, hyperthermia and toxin. One of the defensive response is to instruct the synthesis of a small group of a protein called heat shock proteins. Heat shock proteins (HSPs) are families of molecular chaperone mostly expressed and conserved in all cellular organism and performs several roles in eukaryotic cells. The primary role of HSPs is to maintain and correct the function of protein machinery by maintaining the structure and folding of the client proteins when cells are exposed to homeostatic environment (Chatterjee & Burns, 2017; Li *et al.*, 2009; Swuec & Barlow, 2012; Proia *et al.*, 2015).

The expression of the HSP to different stress factors is known as Heat Shock Response (HSR) regulated at the transcription level by heat shock factors (HSFs) which is the highest transcription regulator of HSPs (Chatterjee & Burns, 2017). HSFs are highly conserved N-terminal binding domain as well as C-terminal transactivation domain and are grouped as HSF1,2,3,4 and HSFY (Chatterjee & Burns, 2017). Of these HSFs, HSF1 mostly regulates HSR as well as HSP gene expression via genes known as heat shock elements (HSE) as shown in Fig. 1. HSF1 works by modulating the expression of HSP. During stress condition, HSF1 is triggered in the cytoplasm by HSP70/90 which binds to HSF1 to block its transcription action. When responding to external stress, HSP separate from to activate HSF1 which binds to HSE by activating HSP genes for promoting cellular protection. Depending on the size and molecular weight HSPs are classified as Hsps 40, 60, 70, 90 and 100. Heat shock protein 90 (Hsp90) is of the most interest and has attracted attention due to its multiple functions (Eachkoti *et al.*, 2014). It is mostly involved in folding and stabilization of numerous protein including those which contribute to the development of cancer (Sauvage *et al.*, 2017). The focus of this thesis is on the Hsp90 as a cancer chemotherapeutic target and is discussed in details in the next subsection.

1.1.2 Structure and function of Hsp90 as cancer chemotherapeutic target

Heat shock protein 90 (Hsp90) is a 90 kDa molecular chaperone highly abundant in stress responses (Proia *et al.*, 2015), with the primary role of controlling cell survival (cell growth and development) (Swuec & Barlow, 2012; Gupta *et al.*, 2015), stabilization and maintaining the active conformations of steroid hormone receptors (Swuec & Barlow, 2012; Jiang *et al.*, 2016;

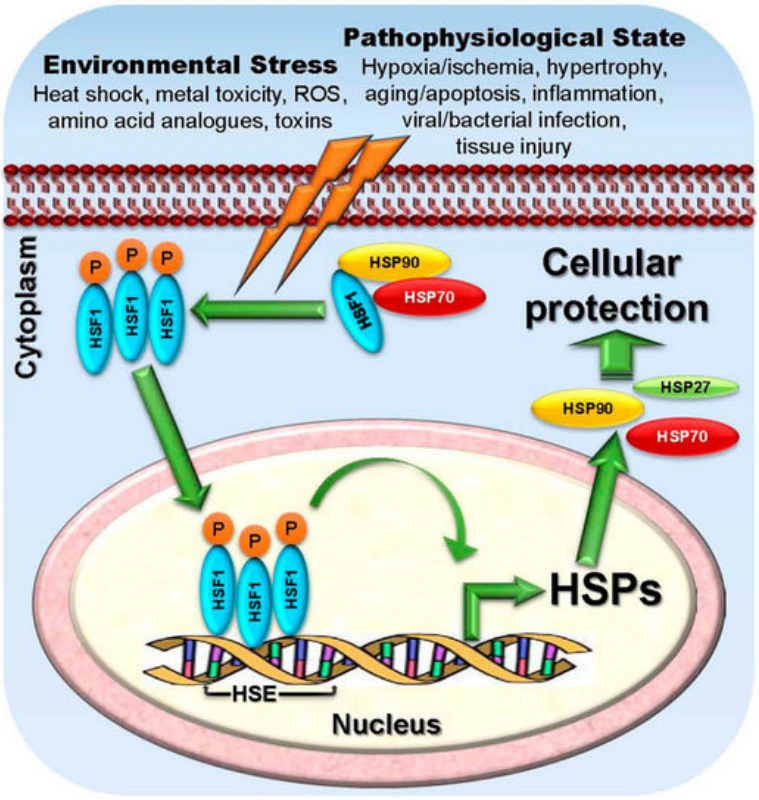
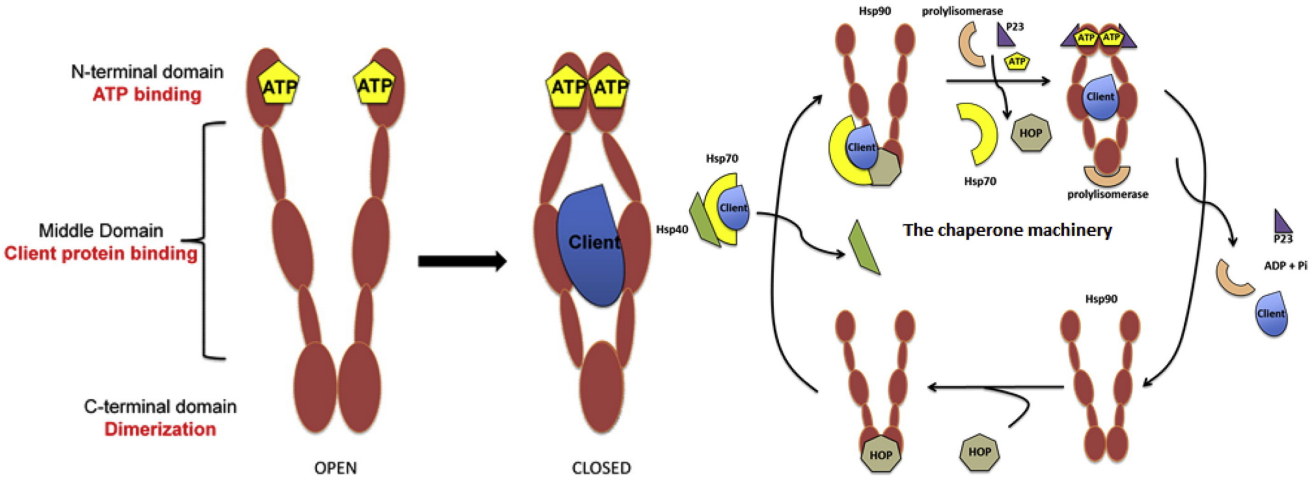


Figure 1: Regulation of HSR and HSP via HSE binding. Adapted from Chatterjee and Burns, (2017).

Yun *et al.*, 2011). It also controls the functions of many signalling proteins such as those involved in pathways of cell proliferation, angiogenesis, cell cycle progression, metastasis and invasion (Li *et al.*, 2009). Over 200 known client proteins expressed in tumour cells depend on Hsp90 for their functions, growth, folding and maturity (Swuec & Barlow, 2012; Proia *et al.*, 2015; Li *et al.*, 2009; Eachkoti *et al.*, 2014; Shi *et al.*, 2012; Sha & Cao, 2015). Such client proteins include; kinases such as CDK, BRAF, AKT, BCR-ABL, CRAF and AGFR. Others are transcription factors such as oestrogen and androgen receptors (AR), p53 and HIF-1 (Proia *et al.*, 2015; Li *et al.*, 2009; Eachkoti *et al.*, 2014; Shi *et al.*, 2012). Thus, Hsp90 has been considered as a potential chemotherapeutic target for the treatment of cancer and other diseases such as alzheimer, parkinson and cardiovascular diseases (Swuec & Barlow, 2012). Inhibition of Hsp90 functions results in the simultaneous depletion of several client proteins (Eachkoti *et al.*, 2014) as well as degradation of oncogenic protein, by stopping multiple signalling transduction pathways of different cancer cells and hence bringing broader anticancer effects (Jiang *et al.*, 2016; Shi *et al.*, 2012).

The structure of Hsp90 (Fig. 2) is composed of three functional domains namely; N-terminal adenosine triphosphate/adenosine diphosphate (ATP/ADP) binding domain, middle domain concerned with client protein binding, and a C-terminal dimerization domain (Proia *et al.*, 2015; Jiang *et al.*, 2016; Shi *et al.*, 2012; Sauvage *et al.*, 2017). The function of the Hsp90 chaperone depends on the domains conformation changes driven by nucleotide binding, hy-

drolysis and exchange of ATP to ADP in the N-terminal domain as well as on the interaction of the client proteins (Sauvage *et al.*, 2017). Several compounds are designed to bind at the ATP-binding site to inhibit the function of Hsp90 (Jiang *et al.*, 2016). In addition, the C-terminal domain has shown a second ATP binding site which needs further exploration (Sauvage *et al.*, 2017).



(a) (b)

Figure 2: (a) The structure of Hsp90 functioning machinery. (b) The changes of the binding client protein affects the conformation of Hsp90 resulting to N-terminal ATP/ADP binding site to open and close. Adapted from Sauvage *et al.* (2017).

An accumulation of evidence showing Hsp90 as a potential cancer target already exists (Barker *et al.*, 2010; Krukenberg *et al.*, 2011; Meyer *et al.*, 2003; Cavenagh *et al.*, 2017). Pharmacological inhibition of Hsp90 results in destabilization and proteasome destruction of client proteins (Sha & Cao, 2015; Gewirth, 2016). Thus, drug candidates are designed to target Hsp90 for the treatment of various cancers and neurodegenerative diseases.

1.1.3 Inhibitors of Hsp90

Several Hsp90 inhibitors (Fig. 3) have been reported over the past years (Roe *et al.*, 1999; Wang *et al.*, 2006; Ge *et al.*, 2006; Tian *et al.*, 2004). Currently, about 20 Hsp90 inhibitors are in clinical trials (Bhat *et al.*, 2014). Two structurally unrelated natural products geldanamycin (**1**), a benzoquinone ansamycin antibiotic and radicicol (**2**) were the first reported Hsp90 inhibitors (Roe *et al.*, 1999). Both compounds showed strong anticancer activities by inducing protein degradation through proteasome but did not progress into clinical trials (Roe *et al.*, 1999). Poor *in vivo* activities, liver toxicity, poor aqueous solubility and difficulties in formulation/production were the clinical limitations for these class of compounds to enter into phase I clinical trial (Soga *et al.*, 2013). Two geldanamycin derivatives; 17-allylamino-demethoxy geldanamycin (17-AAG, tanespimycin (**3**)) and 17-dimethylamino-geldanamycin (17-DMAG, alvespimycin (**4**)) were synthesized and entered into phase I clinical trials (Ge *et al.*, 2006; Tian *et al.*, 2004; Hertlein *et al.*, 2010). In clinical trials, 17-AAG (**3**) showed strong tumour

selectivity and it has been tested as monotherapy or in combination with other agents. Clinical uses of 17-AAG (**3**) were limited due to poor solubility and difficulty in formulation (Ge *et al.*, 2006; Tian *et al.*, 2004; Hertlein *et al.*, 2010). 17-DMAG (**4**) was synthesized to overcome solubility problems of 17-AAG (**3**). Although 17-DMAG (**4**) showed strong Hsp90 inhibition and good aqueous solubility, further progress in clinical trials was hindered by its toxicities (Chiosis, 2006). Classes of compounds with purine scaffolds were then designed and synthesized as Hsp90 inhibitors. Some of these compounds are in preclinical and clinical trials. For instance, BIIB021 (**5**) is in phase I and II clinical trials for patient with metastatic breast cancer. Other purine compounds such as SNX-5422 (**6**) and SN-38 (**7**) were designed and entered into phase I clinical trials. However, they were stopped due to ocular toxicity as well as irreversible retinal damage (Peterson, 2012). None purine compounds, for example, docetaxel (**8**) and gambogic acid (**9**) have shown promising activities. Despite the strong activities shown by these classes of compounds, their clinical uses have been hampered by many factors including toxicities and poor aqueous solubility (Peterson, 2012).

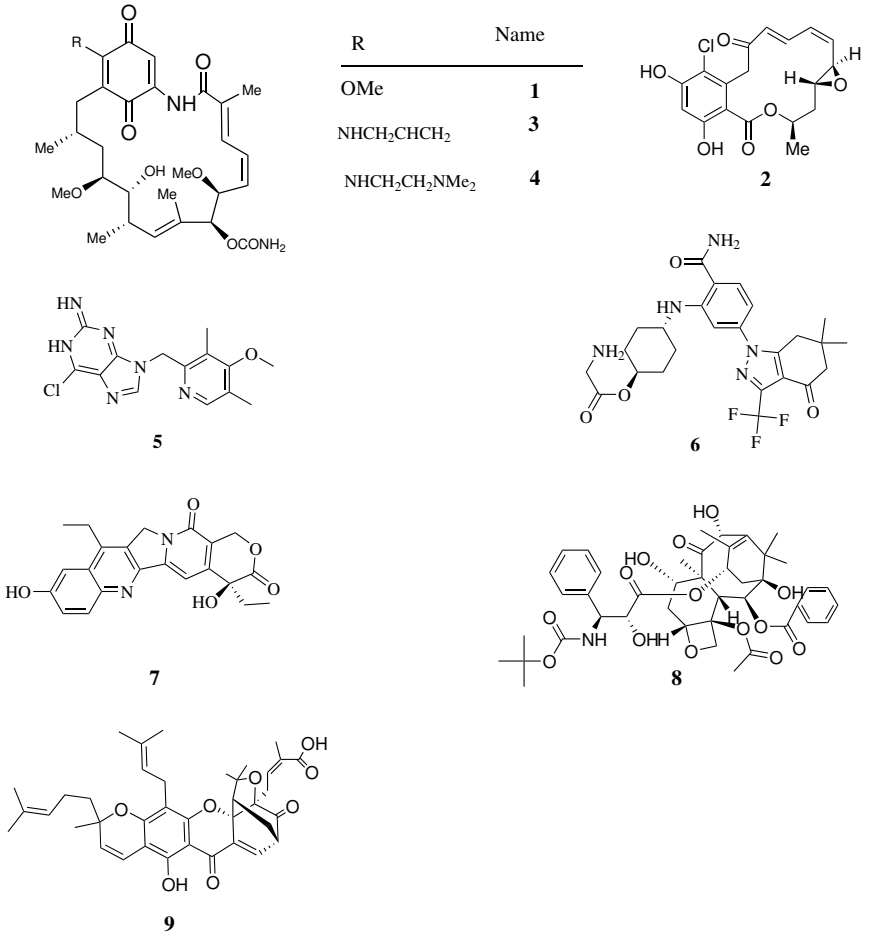


Figure 3: Chemical structures of potential Hsp90 inhibitors clinically limited due to their poor pharmacokinetic.

The major clinical limitation of many drugs is related to their; low efficacy, pharmacologi-

cal toxicity and non-appropriate pharmacokinetic properties of the drug candidate(s) (Puccetti *et al.*, 2005). Such properties were studied during the development phases, while, instead they should have been investigated prior to the development phase. This results in the discovery of related adverse effects and toxicity at a very late stage, a reason for many drugs failure and discontinuity from clinical trials. Until now, there are no approved Hsp90 inhibitors.

Today, advances in computational (*in silico*) methods have provided a means of studying biological activities, pharmacological and pharmacokinetic properties of potential lead compounds before they enter into pre-clinical and clinical trials (Ntie-Kang *et al.*, 2013). Furthermore, computational methods are gaining significant attention in drug repurposing (Sohraby *et al.*, 2019). In particular, molecular docking is a well established and known method in structure based drug design, discovery and development. It predicts the binding free energy of the 3D structure i.e the protein (receptor) and the drug (ligand or small molecule) (De Vivo *et al.*, 2016). Several protocols and tools for performing molecular docking in structure based drug design have been developed. Some protocols for example, consider docking small molecules to the static crystal structure of the protein without water (Kavitha & Velraj, 2017; Qin *et al.*, 2015). It has been shown however that, water plays an important role during protein-ligand interaction (Rarey *et al.*, 1999; Santos *et al.*, 2009; Cui *et al.*, 2018; Huang *et al.*, 2015; Bissantz *et al.*, 2010). Other docking protocols have considered water molecules during molecular docking on the crystal protein structure (Roberts & Mancera, 2008; Thilagavathi & Mancera, 2010; Kumar & Zhang, 2013). Such protocols have shown an increase of docking accuracy when water is included in the binding pocket. Other protocols have showed that, inclusion of water depends on the nature of the ligand, protein system, scoring function as well as the docking methods used (Kumar & Zhang, 2013).

Although incorporation of water into the crystal structure of the protein during molecular docking has shown success and improves the binding energy of some ligands, the approach is complicated by global flexibility of the protein (Roberts & Mancera, 2008). To address this challenge, the relaxed complex scheme (RCS) (Lin *et al.*, 2002) was developed. In this protocol, a molecular dynamics simulation of the protein is performed and after equilibration, receptor configuration are generated and selected for docking. The selected receptor ensembles (snapshots) account for the structure flexibility of the protein for ligand binding (Lin *et al.*, 2002; Roberts & Mancera, 2008). Since the invention of RCS protocol (Lin *et al.*, 2002; Schames *et al.*, 2004) several reports by (Barakat & Tuszynski, 2011; Barakat *et al.*, 2010; Chan *et al.*, 2013) to mention a few, have shown success of using RCS in drug design and discovery. However, the role of water has been neglected in most reports (Barakat & Tuszynski, 2011; Barakat *et al.*, 2010; Chan *et al.*, 2013; Cheng *et al.*, 2008). Reports have shown that water has effects on the binding energy of ligands when docked to different receptor ensembles (Santos *et al.*, 2009). Besides the role of water, other factors that can affect the docking predication is when

one docks ligands to apo or holo conformation (Cheng *et al.*, 2008; Barakat *et al.*, 2010; Michel & Cuchillo, 2012).

In this thesis, different docking protocols for repurposing FDA-approved drugs as Hsp90 inhibitors are investigated. In particular, the effects of inclusion of water in the protein active site of both holo and apo protein structures from the MD simulation are investigated. The finding shows a strong sensitivity to the details of the docking protocol: the use of crystal structure versus an ensemble structures from a simulation, docking with and without water, and finally docking into apo or holo structures.

1.1.4 The role of solvents on host-guest thermodynamics, kinetics and residence time

It is well contented that in drug design, solvents play an important role as they are used in drug formulation, determining the rates of chemical reaction and folding kinetics of molecules (Fu *et al.*, 2014; Papadakis & Deligkiozi, 2019; Levy *et al.*, 2001; da Silva *et al.*, 2018; Sheehan & Sharratt, 1998; Slakman & West, 2019). In many chemical reactions, solvents are known to influence the rate of chemical reaction by controlling the diffusion and rates of the reactants. Difference in solvent dielectric constant, polarizability and polarity are among the factors known to affect the rate of reaction in many biological systems. Exploration of solvent effects on supramolecular interaction (Fu *et al.*, 2014; Papadakis & Deligkiozi, 2019), folding kinetics of biological molecules (Levy *et al.*, 2001) chemical reaction (da Silva *et al.*, 2018; Sheehan & Sharratt, 1998; Slakman & West, 2019) not only has gained much attention in the recent decade but also it has become an active area of research among chemists and chemical physicists. For example, the effect of solvents on thermodynamic stability of supramolecular complexes has been investigated before and found that different polarity in solvents results into different thermodynamic stability in supramolecular recognition (Papadakis & Deligkiozi, 2019). In their work Kang et al., investigated the supramolecular binding of host-guest based on 1-adamantanecarboxylic acid and found that the interaction was entropically favoured, with less polar solvents showing a reverse processes on supramolecular recognition/binding (Kang & Rebek Jr, 1996). Neutral and ionic species such as potassium ions are also involved in supramolecular interactions and solvent effects were also observed for host-guest interactions involving such ions and largely depend on the guest molecule in a given host molecule (Papadakis & Deligkiozi, 2019). As an example, using the linear free energy approach, the host-guest complexation of cyclophane, a hydrophobic molecule and crown ether-18-crown-6 involving potassium ion has been carried out. The association constant of cyclophane was observed to be higher 5-folds in water than in carbon disulfide for a pyrene-cyclophane complex (Smithrud & Diederich, 1990; Izatt *et al.*, 1995). The aforementioned examples, provide an underlying basis on the related solvent effects on kinetics and thermodynamic stability in host-guest interactions. This thesis investigated the solvent related effects on the thermodynamics

and kinetics of host-guest based on chitosan-toussantine-A complex using metadynamics and thermodynamics integration approaches. A detailed literature review on solution conformation of natural products as well as chitosan nanoparticle based drug delivery is given in Chapter 2.

1.2 Statement of the problem

Molecular chaperones are important chemotherapeutic targets in treatments of various diseases. HSPs are chaperones which in recent years have proved to be important targets in the treatment of cancer (Barker *et al.*, 2010; Krukenberg *et al.*, 2011; Meyer *et al.*, 2003; Cavenagh *et al.*, 2017). Of all the HSPs molecular chaperones, Hsp90 has become an important target in cancer treatment. Several efforts to identify potential Hsp90 inhibitors have been attempted. Despite of such efforts, until now there is no any approved Hsp90 inhibitors for cancer treatment. Several approaches, experimental and computational have been employed in investigating Hsp90 inhibitors. Although, experimental methods have provided some useful insights into the problems, further exploration has been limited due to higher degree of similarities of the Hsp90 isoforms (Sha & Cao, 2015). The use of computer simulations can aid experimental approaches in exploration and identification of specific Hsp90 isoform inhibitors. In light of this, several computational protocols such as molecular docking have been employed at large, however, have presented a gap on how such protocols are applied to solve specific problem in drug design. Such protocols have considered thermodynamic information based on static structure of the protein. In addition, the role of water and conformational fluctuation in the Hsp90 has been given less attention and remain unaddressable issue in previous works, such issues are addressed in this thesis. Furthermore, solvent effect on conformation and kinetics of small molecule with their host molecule has been investigated considering thermodynamic data. This gave an opportunity for this study to explore and investigate different docking protocols, the role of water and conformation fluctuation in drug design. The thesis further explored solvent related effects on drugs/natural products conformation and host-guest kinetics and residence time.

1.3 Rationale of the study

The process of drug design, discovery and development is lengthy, costly and a challenging processes. Very recently, the advert of computer hardware and development of algorithms has helped scientists to design and develop drugs in a short time and reduced the related costs. Also, many drugs discovered using traditional drug design approaches are withdrawing from clinical uses due to undesirable pharmacokinetics properties. Such properties were studied during the development phase, instead they should be investigated prior to clinical testing or development. Computer aided drug design approaches have provided a better way to study such properties prior to clinical testing. Owing to the role played by computers in aiding the process of drug

design, this thesis employed various computational methods to study the role of solvents and conformational fluctuations of macromolecules towards drug design.

1.4 Objectives

1.4.1 General objective

The general objective of this thesis is to understand the role played by both solvents and conformational fluctuation of macromolecules and small molecules, which play an important role in determining the biological activities of small molecule under investigation.

1.4.2 Specific objectives

The specific objectives of the present thesis are to:

(i) Compare different docking protocols used in drug design using Hsp90 as a molecular target in cancer treatment.

(ii) Evaluate the solution conformation of drugs/ natural products using curcumin as a model molecule.

(iii) Assess the solvent effects on drug-nanoparticle kinetics and residence time using chitosan-toussantine-A complex as a model system.

1.5 Research questions

(i) What are the sensitivity of different docking protocols used in drug design for the discovery of Hsp90 inhibitors?

(ii) How solvent affects the conformation and orientations of curcumin a model natural product?

(iii) How solvent affects the kinetics and residence time of toussantine-A bound to chitosan nanoparticle?

1.6 Significance of the study

Drug design and discovery is a challenging, costly and lengthy process. The development of computer hardware and algorithms have helped scientists to a great extend to shorten and reduce the related costs. Although, there is an advancement in drug design using computer simulations, several aspects need to be given attention. For example, the effect of solvents and conformation of macromolecules are important factors to be considered. Unfortunately, much efforts have been directed only in understanding the thermodynamics of drug binding, with

little efforts directed on understanding the role of solvents and conformational fluctuations. In addition, drug kinetics and residence time are important factors which have gained less priority. This work has contributed by studying different docking protocols used in drug design, furthermore, the influence of solvents on drug conformation, kinetics and residence time are reported in this thesis. These results are important in the drug design and discovery process. The methods and approaches used in this thesis can further be extended to identify and design other drugs for different indications.

1.7 Delineation of the study

Due to the growing importance of *in silico* drug design which has helped scientists to study molecules at atomistic level. The work reported in this thesis was performed employing different *in silico* approaches *viz* molecular docking, molecular dynamics and enhanced sampling methods such as metadynamics.

CHAPTER TWO

LITERATURE REVIEW

2.1 Molecular docking

In structure based drug design, macromolecules such as protein, DNA and RNA are targets for bioactive agents in the treatment of many diseases. The association of macromolecules and biologically active agents (ligands) has a great role in regulating their functions. Molecular docking is a powerful tool used to find the most favourable binding mode of the ligand when binds to the macromolecule to form a complex. Due to the importance of molecular docking in drug design, over the past decades, several docking programs and algorithms to perform molecular docking have been developed, tested and used in drug design to predict and discriminate between good and bad binders (Ntie-Kang *et al.*, 2013; Sohraby *et al.*, 2019). Such docking programs includes: FLExX (Kramer *et al.*, 1999; Schneider *et al.*, 2013), AUTODOCK/VINA (Trott & Olson, 2010; Morris *et al.*, 1998) MS-DOCK (Sauton *et al.*, 2008), GOLD (Verdonk *et al.*, 2003) and SURFLEX (Jain, 2003) to mention a few. For a detailed and comprehesive review on docking algorithms a reader is refereed to Dias *et al.* (2008). These docking programs are based on different scoring functions and specific search algorithms such as incremental algorithm, genetic algorithm and Monte Carlo. The former can be grouped into three categories: empirical based, knowledge based and force-field based scoring function (Dias *et al.*, 2008).

The scoring functions implemented in many docking programs are able to estimate the binding free energy (ΔG) or binding constant (K_i). The former ΔG is obtained from the Gibbs-Helmholtz as:

$$\Delta G = \Delta H - T\Delta S \tag{2.1}$$

where ΔG is the binding free energy, ΔH is the enthlapy, T is the temperature in Kelvin and ΔS is the entropy. The binding constant (K_i) is related to ΔG as:

$$\Delta G = -RT \ln K_i \tag{2.2}$$

In this study, FlexX part of leadIT was used to introduce the basic concepts of molecular docking. As discussed in the material and methods section, FlexX employs an incremental construction algorithm to bring a ligand into the active site. The ligand is usually decomposed into several fragments, then ligand reconstruction is guided by the protein active site using various placement strategies (Kramer *et al.*, 1999). The binding free energy is calculated and assessed by using the HYDE (HYdrogen bond and DEhydration energy in protein-ligand complex) scoring function (Schneider *et al.*, 2013) which has the form:

$$\Delta G_{\mathrm{HYDE}} = \sum_{\text{atoms } i} (\Delta G^{i}_{\mathrm{dehydration}} + \Delta G^{i}_{\text{H-bonds}}) \tag{2.3}$$

During HYDE calculations, the hydrophobic and hydrophilic dehydration contributions are calculated separately (Schneider *et al.*, 2013) as follows:

$$\Delta G^{i,\,\text{hydrophobic}}_{\mathrm{dehydration}} = -2.3RT \cdot p\log P^{i} \cdot (\mathrm{acc}^{i}_{\mathrm{unbound}} - \mathrm{acc}^{i}_{\mathrm{bound}}) \tag{2.4}$$

$$\Delta G^{i,\text{hydrophilic}}_{\mathrm{dehydration}} = -2.3RT \cdot p\log P^{i} \cdot f^{i}_{\mathrm{bur}} \cdot f^{i}_{\mathrm{water}} \sum_{\text{H-bonds functions } j} w^{j}.P^{j}_{\mathrm{dehydration}} \tag{2.5}$$

f^{i}_{water} is calculated by using equation:

$$f^{i}_{\mathrm{water}} = \sum_{\text{H-bonds functions } j \text{ of atoms } i} \mathrm{water}^{j}_{\mathrm{overlap}} \cdot \mathrm{water}^{j}_{\mathrm{interactions}} \tag{2.6}$$

where; p^{j}_{dehyd} is the probability of dehydration of each hydrogen bond function j (H-bond donor or acceptor); w^{i} is the weight for multiple H-bonds formed by hydrophilic atoms; f^{i}_{bur} is the scaling factor for buriedness of a hydrophilic groups in the unbound state and f^{i}_{water} is the correction factor for local arrangement of water in proximity to H-bond function (Schneider *et al.*, 2013). For docking experiments involving water, the hydration sites of the protein in the active site is determined using the particle concept algorithm (Rarey *et al.*, 1999) implemented in FlexX.

Although molecular docking is faster in calculating the binding free energy, it has limitation related to conformational fluctuation of the macromolecules. Many docking algorithms do not allow the flexibility of the protein. Some algorithms allow only for the side chain flexibility during molecular docking. This limitation has necessitated the rise of new approaches to overcome such a problem. One of such approach is the relaxed complex scheme (RCS) (Amaro *et al.*, 2008b; Schames *et al.*, 2004; Durrant & McCammon, 2011). In RCS, one needs to run MD simulation and extract many ensemble structures which are then subjected to molecular docking programs. Molecular dynamics simulation is performed to provide dynamic and fluctuations of the complexes. Although RCS approach provides flexibility of the protein structure, it is limited by the docking algorithms which require a rigid protein structure (Durrant & McCammon, 2011). In the next section molecular dynamics simulation as a tool in studying not only protein-ligand interactions but also provide dynamical behaviour of the system of interest is discussed.

2.2 Molecular dynamics

Molecular dynamics (MD) simulations are computational simulation methods which enable the movement of molecules or atoms at a nano scale. In a molecular dynamics simulation particles are allowed to interact with each other at a given time, resulting in producing a trajectory which contains information about the dynamical behaviour of the studied system. This section of the thesis starts by discussing the basic ingredients of molecular dynamics. In order to perform a molecular dynamics simulation of a system of interest, one needs to specify the initial conditions i.e velocities and positions of every particle in a system at time zero and interaction potential for deriving the forces among all the particles. Then, the system evolved in time is solved by a set of classical equations of motion for all particles in the system.

In classical mechanics, the force is calculated as the derivative of the interacting potential, V(R), and the force acting on each particle in the system is determined using the Equation 2.7.

$$F_i = m_i a_i = m_i \frac{d^2 r_i}{dt^2} = -\nabla V(R) \tag{2.7}$$

Equation 2.7 provides a set of 3N coupled second order differential Equations that can be propagated forward and backward in time. In Equation 2.7, F_i represents a vectorial sum of all forces acting on i^{th} particle of the system. The forces are derived from a potential energy function $V(r^N)$ that describes how particles in a given system interact to each other. Equation 2.7 is solved by integration using a small time steps in an order of magnitude ranging between 1 to 10 fs. This implies that at every time step of the simulation, both positions and velocities of every single atom are calculated. The initial positions and velocity need to be known, usually specified in an input datafile, and the forces are calculated through the potential energy. The initial coordinates at t = 0 is obtained from experimental data while the velocities are not known and they need to be assigned. Assigning of velocity can be done using a Maxwell-Boltzmann distribution at a given temperature, usually at 300 K for many biological systems.

2.2.1 Energy minimization

Since the starting configurations are obtained from experimental data, usually in their crystallographic structures they are far from equilibrium. Thus, an energy minimization is needed to be done before performing MD simulation. Energy minimization is done to reduce the excessively large interaction forces within atoms, and therefore, bring the system at a relaxed state. Algorithms such as steepest descent (SD), conjugate gradient (CG) and Newton's methods are used to do energy minimization.

2.2.2 Force fields

Due to the current limitation of computational resources, it is not possible to accurately calculate the potential energy of macromolecules by using quantum chemistry electronic approaches. Thus, force fields which are simple classical functions with a number of parameters from either experimental data or quantum chemistry calculation are used. A force field is, therefore, a set of parameters containing specific function which describes the interaction of atoms. Generally, a force field contains functional terms such as Lennard-Jones, coulomb, harmonic vibration and model parameters which describe the function terms. Several families of force field are available for molecular simulations, for example, AMBER (Lindorff-Larsen *et al.*, 2010; Wang *et al.*, 2004), CHARMM27 (MacKerell Jr *et al.*, 2000), GROMOS (Oostenbrink *et al.*, 2004) and OPLS (Jorgensen & Tirado-Rives, 1988; Jorgensen *et al.*, 1996). Since these force fields are used for different systems, one has to choose the force field depending on the system to be simulated. In this thesis AMBER03 and OPLS-AA force fields are used and described in details, in particular, the OPLS (Optimized Potential for Liquid Systems) force field which has a potential function form described in Equation 2.8.

$$\begin{aligned} V(r^N) = & \sum_{\text{bonds}} k_b (l - l_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 \\ & + \sum_{\text{torsion}} \left[\frac{V_1}{2}(1 + cos(\phi)) + \frac{V_2}{2}(1 - cos(2\phi)) \frac{V_3}{2}(1 + cos(3\phi)) \frac{V_4}{2}[1 - cos(4\phi)] \right] \\ & + \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left[V_{ij}^{\text{Coulomb}}(r_{ij}) + V_{ij}^{LJ}(r_{ij}) \right] \end{aligned} \tag{2.8}$$

In Equation 2.8 each term is an approximation for different interaction energy, and the first three terms represent covalently bonded terms linked by bonds (l), angles (θ) and torsions (ϕ), while the last term represents non-bonded term for long-range electrostatic and van der Waals energies (Lennard-Jones 6-12 potential) (Jorgensen & Tirado-Rives, 1988; Jorgensen *et al.*, 1996).

2.2.3 Integrating the equation of motion

Once the forces for each atom are obtained from the force field, both velocity and position at every particle at $t = t_0 + \Delta t$ can now be computed by integrating Equation 2.7. In general, the time step Δt, is a small number with the order of ($1fs = 10^{-15}$s), the use of longer time steps usually affect the simulation and conservation of energy.

Several algorithms such as Verlet (Verlet, 1967), velocity Verlet (Swope *et al.*, 1982), Leap frog (Van Gunsteren & Berendsen, 1988) are available and used to integrate the Newtons equation of motion. Here, the Verlet algorithm which is often used to integrate the equation of motion

is described (Verlet, 1967). This algorithm is derived by the Taylor expansion of coordinate of the particle at time $t \pm \Delta t$, where Δt is the time step.

$$\vec{r}_i(t \pm \Delta t) = \vec{r}_i(t) \pm \vec{v}_i(t)\Delta t + \frac{\vec{a}_i(t)}{2!}\Delta t^2 \pm \frac{\dddot{\vec{r}}_i}{3!}\Delta t^3 + O(\Delta t^4) \tag{2.9}$$

summing up gives;

$$\begin{aligned} \vec{r}_i(t+\Delta t) &= 2\vec{r}_i(t) - \vec{r}_i(t-\Delta t) + \vec{a}_i(t)\Delta t^2 + O(\Delta t^4) \\ \vec{r}_i(t+\Delta t) &\approx 2\vec{r}_i(t) - \vec{r}_i(t-\Delta t) + \vec{a}_i(t)\Delta t^2. \end{aligned} \tag{2.10}$$

Equation 2.10 can be implemented in classical molecular dynamics to propagate the positions of each atom at a time. In this Equation 2.10 the estimation of new position contains an error in the order of Δt^4. However, this equation does not explicitly include the velocity terms. The required velocity is constructed from the trajectory using Equation 2.11.

$$v_i(t) = \frac{r_i(t+\Delta t) - r_i(t-\Delta t)}{2\Delta t}. \tag{2.11}$$

2.3 Metadynamics as an enhanced sampling method

Despite the fact that MD simulation is widely used in many aspects, it suffers from sampling problem and time scale, as it does not allow escape from local free energy minimum and requires long time for the system to reach ergodicity (Laio & Parrinello, 2002; Barducci *et al.*, 2008). To overcome such limitations, several algorithms in the name of enhanced sampling methods such as metadynamics (Laio & Parrinello, 2002; Barducci *et al.*, 2008), umbrella sampling (Kästner, 2011) and conformational flooding (Grubmüller, 1995) to name a few have been introduced.

2.3.1 Theory of metadynamics

Metadynamics is one of the enhanced sampling methods used to sample rare events (Laio & Parrinello, 2002). In metadynamics, the added bias potential discourages the system from revisiting already visited phase space allowing one to surmount energy barriers. Figure 4 provides a picture on how metadynamics work. Consider a system of N-particles interacting with the potential $V(r)$ in a canonical ensemble with a temperature T. In a given system, if a dynamics stucks at one local energy minimum with $V(r)$, the probability to escape to other local minimum is low. In metadynamics, a history-dependent bias potential, $V(r)$, constructed along the selected reaction coordinate $\vec{s}(q)$, called collective variables (CVs) is added to the Hamiltonian of the system as follows:

$$H = T + U + V(r) \tag{2.12}$$

where 2.12, T is the total kinetic energy, U is the potential energy which describe the system and $V(r)$ is the bias potential. This potential is built as a sum of Gaussian kernels (height (W) and width (δ_i)) deposited along the trajectory in the CVs space, and can be written as:

$$V(r)(\vec{s},t) = \sum_{k\tau<t} W(k\tau) \quad \exp\left(-\sum_{i=1}^{d} \frac{(S_i - s_i(q(k\tau)))^2}{2\delta_i^2} \right) \tag{2.13}$$

where t is the deposition time and τ is the time interval where Gaussian potential with height is added on the position s_i (q(k τ)) of the biased molecules.

Figure 4, describes how metadynamics work for a simple 1D potential with two minima, i.e one for reactant and one for product. In (A) the reaction is at local minimum, then the bias potential is added during the simulation as shown in (B) to discourage the system to revisit already visited phase space, at this point the bias potential is sampling around the minimum. In (C) the added bias potential is large enough forcing the system (in blue) to move to another global energy minimum on the right side. In (D) the added bias potential becomes large than the free energy barrier and the system can easily move from one minimum to another.

Although metadynamics has several benefits, its implementation is faced up by two disadvantages. Firstly, the choice of collective variable (CVs) is nontrivial and a complicated process, as there is no systematic way of selecting appropriate CVs. The Gaussian sigma are chosen from a short standard MD simulation which can provide an overview of the system. The other disadvantage is the difficult in bias potential to converge the underlying potential energy due to the constant height of Gaussian hills.

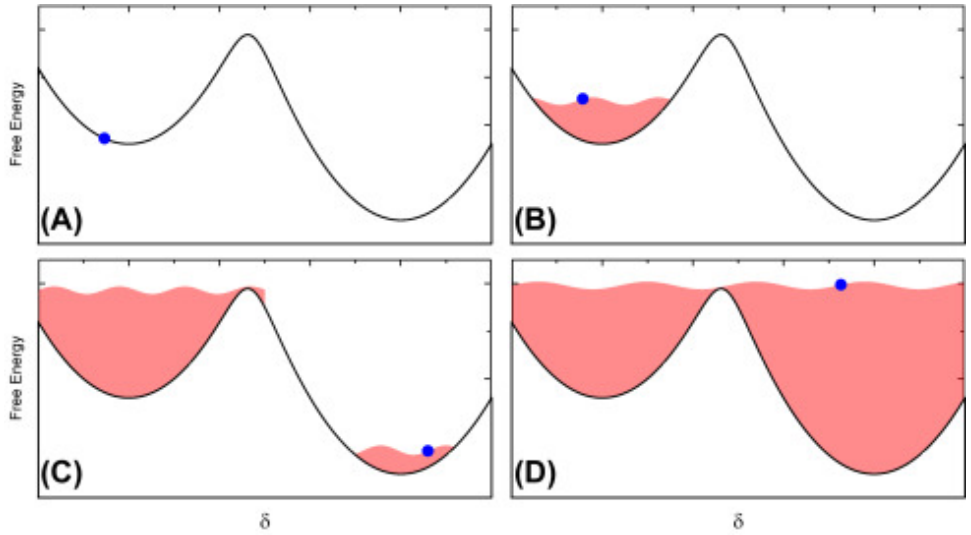


Figure 4: An illustration of how metadynamics work. The blue point represents the position of the system in an arbitrary free energy along the reaction coordinate δ (CVs). Adapted from Jambrina and Aldegunde (2016).

In order to control errors and convergence during metadynamics simulation, well-tempered metadynamics (WT-MetaD), which is a variant of metadynamics was introduced (Barducci *et al.*, 2008). In WT-MetaD the Gaussian height (W) decreases with simulation time as shown in Equation. 2.14.

$$W(k\tau) = W_0 \quad \exp\left(-\frac{V(\vec{s}(q(k\tau)),k\tau)}{k_{\mathrm{B}}\Delta T}\right) \tag{2.14}$$

where, W_0 is the initial Gaussian height, k_B is the Boltzmann constant, ΔT is the input parameter with dimension of temperature that control how the W reduces as the well is filled. To obtain the best efficiency, the parameters W_0 and ΔT are chosen, and the free energy is computed using Equation 2.15

$$F(\vec{s}) = -\frac{T+\Delta T}{\Delta T}V(\vec{s},t\to\infty) \tag{2.15}$$

where, ΔT is the input parameter with dimension of temperature, T is the temperature of the system. In well-tempered metadynamics the bias factor γ is defined as the ratio between the CVs temperature (T+ΔT) as indicated in Equation 2.16

$$\gamma = \frac{T+\Delta T}{T} \tag{2.16}$$

2.3.2 Reweighting free energy

During metadynamics simulation, the free energy is calculated using the biased CVs, the free energies of the unbised CVs can be calculated by reweighting approaches. The probability distribution of the biased CVs from the trajectory is obtained as in Equation 2.17. The quantities on the right hand side represent the probability of sampling a vector at position, q and momentum, p.

$$P(p,q) \propto \exp\left(-\frac{F(\overrightarrow{s})}{k_{\mathrm{B}}T}\right) \tag{2.17}$$

Very recently, a simple reweighting algorithm to recover distribution of unbiased CVs was developed by Tiwary and Perrinello (Tiwary & Parrinello, 2014). Based on this algorithm, by using the internal potential $U(\overrightarrow{r},t)$ and the bias potential $V(\overrightarrow{s},t)$, the probability distribution $P(\overrightarrow{r},t)$ can be written as:

$$P(\overrightarrow{r},t) = \frac{\exp\left(-\beta(U(\overrightarrow{r},t)+V(s(\overrightarrow{r},t))\right)}{\int d\overrightarrow{r}\exp\left(-\beta(U(\overrightarrow{r},t)+V(s(\overrightarrow{r},t))\right)} \tag{2.18}$$

Equation 2.18 can be rewritten in the following form:

$$P(\overrightarrow{r},t) = \exp\left(-\beta(V(s(\overrightarrow{r},t)-c(t))\cdot P_o(\overrightarrow{r})\right) \tag{2.19}$$

where $P_o(\overrightarrow{r})$ is the probability distribution which is not biased, $c(t)$ is time dependent function which gives important information of the system and is defined as:

$$c(t) = \frac{1}{\beta}\log\frac{\int d\overrightarrow{s}\exp\left(-\beta F(\overrightarrow{s})\right)}{\int d\overrightarrow{s}\exp\left(-\beta F(\overrightarrow{s})\right)+V(\overrightarrow{s},t)} \tag{2.20}$$

During WT-MetaD, $c(t)$ is calculated as follows:

$$\exp\left(\beta c(t)\right) \approx \frac{k_{\mathrm{B}}\Delta T}{\gamma\omega\Delta t(2\pi)^{D/2}\mathrm{det}\sigma} = \int d\overrightarrow{s}\left[\exp\left(\gamma V(\overrightarrow{s},t+\Delta t)/k_{\mathrm{B}}T\right)-\exp\left(\gamma V(\overrightarrow{s},t)/k_{\mathrm{B}}T\right)\right] \tag{2.21}$$

In Equation 2.21, $\mathrm{det}\sigma$ is the determinant of the collective variable space and Δt is the time interval between Gaussian depositions.

2.4 Binding free energy in drug design and discovery

It has been mentioned that accurate predication of the affinity of drug to its target receptor (protein or nanocarrier), is one of the *holy grails* in computational drug design (Aldeghi *et al.*, 2016; Ytreberg *et al.*, 2006). Free energy calculations (ΔF) are of particular interest, and have a variety of applications in drug design (Aldeghi *et al.*, 2016; Ytreberg *et al.*, 2006), such as protein-ligand interaction (Aldeghi *et al.*, 2016; Ytreberg *et al.*, 2006; Peräkylä & Nordman, 2001) and solubility of small molecules. Free energy calculation has raised interest to both experimental and computational/theoretical communities (Ytreberg *et al.*, 2006; Ytreberg *et al.*, 2006). Due to its importance in the field of drug design, several methods have been developed. Exisiting free energy methods are classified as equilibrium and nonequilibrium. Equilibrium free energy are those which requires or relies on fully sampled equilibrium simulation done in each step of the free energy calculation. Such methods includes free energy pertubation (FEP) (Rao *et al.*, 1987), MM-PB(GB)SA (Kollman *et al.*, 2000), thermodynamics integration

(Kirkwood, 1935), weighted histogram analysis (Kumar *et al.*, 1992) and Bennett analysis (Shirts & Pande, 2005). In equilibrium free energy, one needs to attain the equilibration at each stage to avoid obtaining biased results. On the other hand, nonequilibrium free energies based on Jarzynski's equality have recently been used in many molecular systems, however they suffer from bias. This work describes equilibrium free energy which is central of this work. Readers interested in nonequilibrium free energy are refereed to (Jarzynski, 1997).

2.4.1 Thermodynamic integration

Thermodynamic integration (TI) is one of the mostly reported accurate alchemical method used to calculate/compare the free energy of two given states A and B (see Fig. 5). The two given states have their potential energies U_A and U_B with different dependence on coordinates. In TI, the free energy is calculated using a thermodynamic path from state A to B by integrating over the ensemble average along the path. As shown in Fig. 5, the potential energy U_A and U_B is calculated as an ensemble average over the sampled phase space from molecular dynamics (MD) or Monte Carlo (MC) simulation. A new potential function is defined as:

$$U(\lambda) = U_A + \lambda(U_B - U_A) \tag{2.22}$$

Here, λ, is a coupling paramenter with the value ranging from 0 to 1. The potential energy as a function of λ varies from the energy of system A for λ = 0 and system B for λ = 1. Taking an example for canonical ensemble, the partition function of the system can be writen as follows:

$$Q(N,V,T,\lambda) = \sum_s \exp\left[\frac{-U_s(\lambda)}{k_\mathrm{B}T}\right] \tag{2.23}$$

$U_s(\lambda)$ is the potential energy of state S in the ensemble function $U(\lambda)$. The corresponding free energy of the system can be defined as follows,

$$\Delta G(N,V,T,\lambda) = -k_\mathrm{B}T\ln Q(N,V,T,\lambda) \tag{2.24}$$

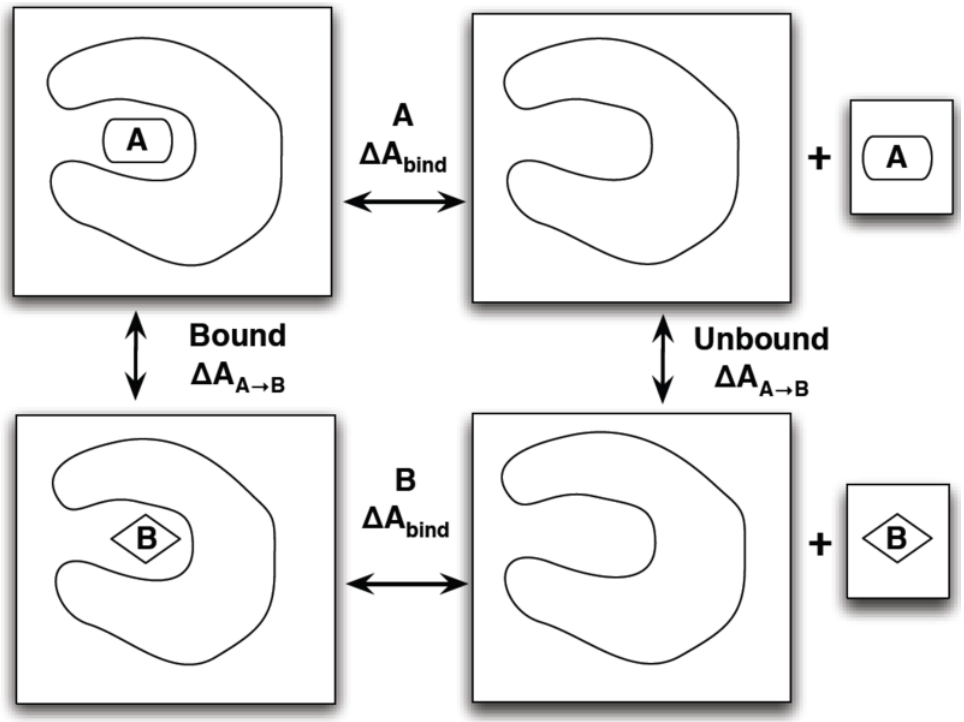


Figure 5: Thermodynamic cycle. Figure adapted from alchemical website at www.alchemistry.org/wiki/Thermodynamic_cycle.

Taking the derivative of G with respect to λ, one can get an ensemble average of the derivative of potential energy with respect to λ.

$$\Delta G(A \to B) = \int_0^1 \frac{\delta F(\lambda)}{\delta \lambda} d\lambda = -\int_0^1 \frac{k_B T}{Q} \frac{\delta Q}{\delta \lambda} d\lambda \tag{2.25}$$

$$= \int_0^1 \frac{k_B T}{Q} \sum_s \frac{1}{k_B T} \exp\left[\frac{-U_s(\lambda)}{k_B T}\right] \frac{\delta U_s(\lambda)}{\delta \lambda} d\lambda \tag{2.26}$$

$$= \int_0^1 \left\langle \frac{\delta U(\lambda)}{\delta \lambda} \right\rangle_\lambda d\lambda \tag{2.27}$$

$$= \int_0^1 \left\langle U_B(\lambda) - U_A(\lambda) \right\rangle_\lambda d\lambda \tag{2.28}$$

Thus, the free energy can be computed from the integral of the ensemble average derivatives of potential energy using a compling parameter, λ. Generally, TI is most accurate method of estimating the binding free energy as it needs extensive sampling of both complex and ligand in solution and the unphysical intermediate states.

2.4.2 Molecular Mechanics Poisson-Boltzmann (Generalized Born) Surface Area (MM-PB(GB)SA)

In structure-based drug design (SBDD), one of the important goal is to identify new small molecules (ligands) which can bind to their macromolecules receptors. The binding of a ligand

to their receptors is described in the following chemical reaction:

$$R + L \rightarrow R : L \tag{2.29}$$

where, R denotes receptor usually protein, DNA or any macromolecule, L denotes a ligand. The strength of the ligand binding to macromolecules is determined by calculating the binding free energy, ΔG_{bind}. Normally, this free energy is determined using experimental approaches, which again suffer from the limitation of high cost and time consuming. Thus, several computational methods have been developed with the aim of reducing both cost and time while achieving the goal of drug design.

MM-PB(GB)SA is the most popular method used to estimate the binding free energy, ΔG_{bind} of ligands to their macromolecules in drug design. Although, MM-PBSA depends on molecular dynamics simulation of the complex, it does not demand high computational cost unlike the alchemical pertubation (AP) approaches. However, the accuracy of MM-PBSA lies between the scoring and AP approaches. Because of less computational demand and the accuracy, today, MM-PBSA is used in many systems to estimate the binding affinity of ligands with different degree of success. In MM-PBSA, the ΔG_{bind} from Equation 2.29 is obtained by estimating the free energy of reactant and product.

$$\Delta G_{\text{bind}} = \langle G_{RL} \rangle - \langle G_R \rangle - \langle G_L \rangle \tag{2.30}$$

This method in Equation 2.30 was first reported by (Kollman *et al.*, 2000) and it has significantly gained attention in many application ranging from protein-ligand interaction (Aldeghi *et al.*, 2016; Ytreberg *et al.*, 2006), protein-protein interaction (Chen *et al.*, 2016) as well as conformational stability (Brice & Dominy, 2011). The free energy of each state RL, R and L in equation 2.30 is estimated as the sum of the following components (Kollman *et al.*, 2000).

$$G = E_{\text{bonded}} + E_{\text{ele}} + E_{\text{vdW}} + G_{\text{polar}} + G_{\text{non-polar}} - TS \tag{2.31}$$

(i) Molecular mechanics (MM) energy terms

The first three terms in Equation 2.31 are molecular mechanics (MM) energy terms from bonded, electrostatic and van der Waals interactions. The next two terms, that is, polar and non-polar contribute to solvation free energies. The former is obtained by solving the Poisson-Boltzman (PB) or Generalized Born (GB) equation giving the term MM-PBSA or MM-GBSA. The non-polar contributions is estimated by linearly relating to the solvate accessible surface area (SASA) model. The last term is the absolute temperature T multiplied by entropy S contributions (Genheden & Ryde, 2015). Polar and non-polar contribution to binding energy are further detailed in the next subsections.

(ii) Polar and non-polar solvation energy terms

The free energy of solvation, ΔG_{solv} is calculated from two components, polar and non-polar solvation energies (Genheden & Ryde, 2015) (see Equation 2.32). Polar solvation can be calculated by solving the PB linearly (Equation 2.33) when both ionic strength and solvent potential are low and when asymmetric electrolytes are considered (Genheden & Ryde, 2015).

$$\Delta G_{\text{solv}} = \Delta G_{\text{solv, polar}} + \Delta G_{\text{solv, non-polar}} \tag{2.32}$$

$$\nabla.\varepsilon\nabla\phi = -4\pi\rho_0 + \varepsilon_v K^2\phi \tag{2.33}$$

where, $K^2 = \frac{8\pi e^2 I}{\varepsilon_v k_B T}$, V = solvent, I = ionic strength of the solution, it is defined as $I = Z^2C$, where, Z^2 and C are charge and molar concentration of ion, respectively. When the solvent potential, ϕ, is solved, polar solvation can be computed as follows:

$$\Delta G_{\text{polar}} = \frac{1}{2}\sum q_i\phi_i \tag{2.34}$$

The polar solvation term stated in Equation 2.31, was first developed by numerically solving PB equation (Kollman *et al.*, 2000). However, methods such as continuum-solvation are now commonly used in calculating polar solvation in GB giving the name MM-GBSA (Genheden & Ryde, 2015). Several works exist where the performance of both MM-PBSA and MM-GBSA methods have been compared (Xu *et al.*, 2013; Sun *et al.*, 2014) with MM-PBSA showing better performance (Xu *et al.*, 2013) or similar results (Sun *et al.*, 2014; Rastelli *et al.*, 2010; Sun *et al.*, 2014) depending on the system studied. The effects of polar solvation on PB results have been investigated. For example, it has been shown that the contribution of polar solvation energy to the binding free energy in PB method is affected by the radii used. Other studies have investigated other solvation methods such as GB and polarized continuum model (PCM), and have found that for ligands with same net charge all methods (PB, GB, PCM) gave similar results within 2-5 kJ/mol (Genheden & Ryde, 2015).

The non-polar solvation energy is directly related to the SASA (Genheden & Ryde, 2015; Sun *et al.*, 2014) using the Equation 2.35

$$\Delta G_{\text{non-polar}} = \gamma \times SASA + b \tag{2.35}$$

Here, γ is the coefficient related to surface tension of the solvent and it varies from 0-4.2 kJ/mol (Genheden & Ryde, 2015), b is the fitting parameter.

(iii) Electrostatic term, E_{ele}

In MM-PBSA, electrostatic term, E_{ele}, is calculated using dielectric constant, ε. Originally, MM-PBSA was calculated using a ε =1, although a large value has been suggested for improving the results (Genheden & Ryde, 2015). However, the optimal ε value has been observed to depend on the nature of the binding pocket where a more charged pocket requires a large ε value compared to a more hydrophobic pocket. The ε value ranging from 2-4 has been shown to produce good results in many studies (Genheden & Ryde, 2015).

2.4.3 Limitation of MMPBSA

Despite of its popularity in estimating free energy of ligands bound to their macromolecules, MM-PBSA method is limited due to poor precision. The method results into large standard error deviation (Genheden & Ryde, 2015; Peräkylä & Nordman, 2001). Poor precision renders the method useless especially when one wants to compare results obtained from different approach or when one compares the ligands with similar affinity (Genheden & Ryde, 2015). Today, several suggestions have been attempted to reduce such limitation, one of the attempt has been suggested to run many short independent simulation instead of a single long simulation which can underestimate uncertainty in results (Genheden & Ryde, 2015). Some studies have shown that, an equilibration of 100 ps and production of 100-200 ps of many independent simulation i.e 20-50 ns can reduce the error and would be appropriate enough (Genheden & Ryde, 2015).

2.5 Some examples of successful applications of molecular dynamics and related methods in drug discovery and design

2.5.1 The success and limitations of molecular docking

The preceding sections have discussed several computational methods used in drug discovery and design. This section, discusses the practical application of molecular dynamics and related methods in drug discovery and design. The applications of these method have brought great success in drug design as some drugs discovered using these methods are now available in the market and have been approved by relevant authorities. In this thesis the discussion focuses more on the discovery of the Hsp90 inhibitors. The structure and function of the Hsp90 is described in Chapter 1.

As described in Section 2.1, molecular docking is a popular computational chemistry shape based method with high ability to predict the binding mode of ligands to their 3D macromolecular active sites (Piaz *et al.*, 2012; Abbasi *et al.*, 2018; Saxena *et al.*, 2010; Verdonk *et al.*, 2003; Stjernschantz & Oostenbrink, 2010; Vijesh *et al.*, 2013; Mbatha, 2015; Ramírez & Caballero, 2016), but, it poorly estimates and ranks the binding affinity of ligand to their macromolecules

(Shamsara, 2018; Warren *et al.*, 2006). Docking is widely used in identification of protein binding pockets (Ruppert *et al.*, 1997), drug discovery (lead optimization) (Piaz *et al.*, 2012; Vijesh *et al.*, 2013), protein-protein interaction (Dominguez *et al.*, 2003; Comeau *et al.*, 2004; Tovchigrechko & Vakser, 2006) as well as in virtual screening (Verdonk *et al.*, 2004; Cavasotto *et al.*, 2007) where a large number of small molecules are screened against a known 3D macromolecular target. Other areas where molecular docking is widely used include protein engineering and enzymatic reaction mechanism (Mbatha, 2015). Molecular docking mostly helps to understand the binding nature or orientation of ligand in the protein active site (Mbatha, 2015). During molecular docking, the receptor-ligand binding energy is analyzed using molecular mechanics (Mbatha, 2015) in Equation 2.36.

$$E_{\text{binding}} = E_{\text{target-ligand}} - (E_{\text{target}} + E_{\text{ligand}}) \tag{2.36}$$

Results obtained from molecular docking are either supplemented with experimental studies (Piaz *et al.*, 2012; Gupta *et al.*, 2015), molecular dynamics (Abbasi *et al.*, 2018; Sha & Cao, 2015) and /or binding free energy calculation (Sha & Cao, 2015) to understand the stability of the complex and estimate the binding affinity, respectively. Thus, several studies (Piaz *et al.*, 2012; Abbasi *et al.*, 2018; Saxena *et al.*, 2010) have employed molecular docking to understand the binding modes of ligand inside the Hsp90.

The work by Piaz et al., (2012) investigated the inhibition of Hsp90 by a natural iminosugar (+)-Lentiginosine (**10**) using biochemical methods and molecular docking. Molecular docking revealed that iminosugar (+)-Lentiginosine binds to the middle domain of the Hsp90 differently from the known ATP binding site (Piaz *et al.*, 2012). Furthermore, Abbas *et al.* (2018) carried out a combinations of pharmacophore modeling, molecular docking and molecular dynamics to predict new inhibitors of Hsp90 based on isoxazole scaffold (Abbasi *et al.*, 2018). Molecular docking was performed on 16 training set to provide insight into their binding modes and interaction with Hsp90. Of the 16 compounds, seven showed stability during MD simulation by forming hydrogen bonds with amino acid Asp93 and Thr184 (Abbasi *et al.*, 2018). In a similar study Saxena *et al.* (2010), carried out pharmacophore modeling and molecular docking study to identify new inhibitors of heat shock protein 90. Molecular docking helped to establish the binding mode as well as the amino acids involved in Hsp90-ligand interaction (Saxena *et al.*, 2010). Docking estimated the binding affinity with a correlation coefficient of 0.699 to the experimental IC_{50} values. It was established that, hydrogen bonding and hydrophobic interactions were essential in Hsp90-ligand interaction (Saxena *et al.*, 2010). Molecular docking and pharmacophore modeling both predicated compound 5-(2,4-Dihydroxy-5-isopropyl-phenyl)-4-(4-piperidin-1-ylmethyl-phenyl)-isoxazole-3-carboxylic acid ethyl amide (**11**) to be the most potent inhibitor of Hsp90.

Sepehri and Ghavami (2018) conducted an *in silico* study employing molecular docking and comparative molecular field analysis (CoMFA) to design new inhibitors of Hsp90 based on tetrahydropyrido [4, 3-d] pyrimidine derivatives. In their study, three new molecules were designed and all were docked to the ND-Hsp90 (Sepehri & Ghavami, 2018). Molecular docking helped in establishing the binding mode of the compounds into the ND-Hsp90. Hydroxyl group on phenyl ring was observed to be important in forming hydrogen bonds with hydrophilic residues and conserved water (Sepehri & Ghavami, 2018). The compounds were observed to interact with Phe138 and Asp54 and Asp51 by forming π-sigma and hydrogen bonds (Sepehri & Ghavami, 2018).

In some cases, molecular docking results have shown high correlation with experimental results. For instance, in an effort to discover and develop new potent inhibitors of Hsp90, Gutpa et al., carried out molecular docking of mannich base against Hsp90 and then synthesized the highly ranked compounds (Gupta *et al.*, 2015). Synthesized compounds were then tested experimentally for their anticancer and Hsp90 inhibition. Results showed that activity profile of the synthesized mannich base compounds derived from 2,4-dihydroxyacetophenone/5-chloro2,4-dihydroxyacetophenon correlated well with the docking results (Gupta *et al.*, 2015). Results from molecular docking showed that water molecule and amino acid residues Leu48, Lys58 and Asn106 were necessary in forming hydrogen bonds contact, while amino acid residues Phe138, Val150 and Val186 were important for hydrophobic interactions (Gupta *et al.*, 2015). In another work, Gupta *et al.,* conducted molecular docking and experimental inhibition of Hsp90 using schiff bases derived from 2,4-dihydroxy benzaldehyde/5-chloro-2,4-dihydroxy benzaldehyde (Gupta *et al.*, 2014). Compounds which showed lower binding energies were synthesized and tested experimentally for their Hsp90 inhibition and anticancer activity using MTT (3-(4,5-dimethythiazol-2-yl)-2,5-diphenyltetrazoliumbromide) assays (Gupta *et al.*, 2014). Docking further showed that water molecule 903 and amino acid residue Asp93 and Asn51 were crucial in forming hydrogen bonds (Gupta *et al.*, 2014). Findings reported by Gupta *et al.* (2014) and Gupta *et al.* (2015) shows that both mannich and schiff bases could be developed as potent inhibitors of Hsp90 for cancer treatment.

Pan inhibition of Hsp90 inhibitors has remained the greatest challenge in the discovery and development of potent inhibitor for this target. Several proposed inhibitors of Hsp90 inhibit all the Hsp90 isoforms and hence present the greatest challenge. Pan inhibition of Hsp90 has been associated with toxic side effects of many Hsp90 inhibitors due to the inhibition of hERG activities (Peterson, 2012). Thus, in order to identify specific inhibitors of Hsp90 isoform, Sha and Cao (Sha & Cao, 2015), carried out molecular docking, MD simulation and binding free energy study of a benzolactam compound (**12**) against Hsp90 isoforms (Hsp90β/α and Grp94) to further provide basis on isoform selectivity inhibition. The benzolactam compound was previously reported experimentally (Ernst *et al.*, 2014) to possess isoform selectivity inhibition for

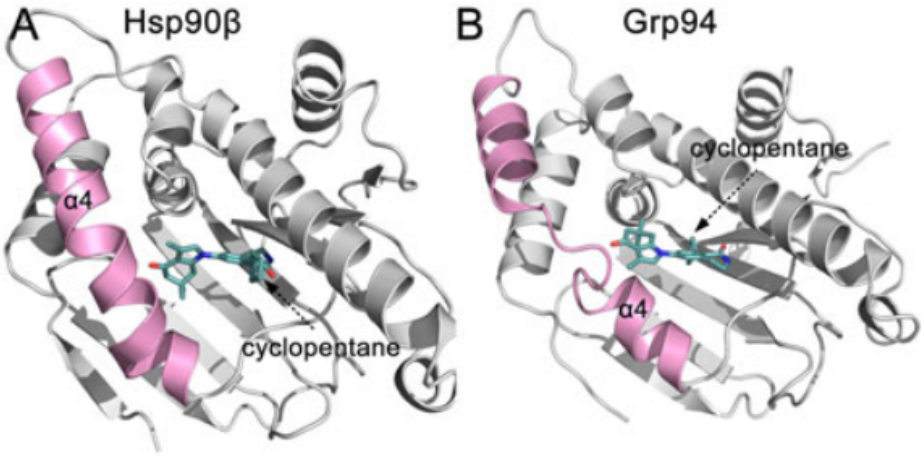


Figure 6: Structure and binding mode of benzolactam (compound **12**) when docked to Hsp90β and Grp94. The figure is reproduced from Sha & Cao, 2015.

Hsp90β/α by >1000-folds against Grp94, however, the structural basis of its selective inhibition remained not well addressed (Sha & Cao, 2015). Molecular docking provided detailed information on its selective inhibition. Compound (**12**) bound differently in Hsp90 and Grp94 (Sha & Cao, 2015). The binding mode of cyclopentane moiety inside Hsp90 and Grp94 is shown in Fig. 6. In both Hsp90β/α, the lactam moiety of compound (**12**) formed hydrogen bond with amino acids Thr1834 and Asp93, respectively (Sha & Cao, 2015). Furthermore, the cyclopentane moiety was buried inside the hydrophobic pocket formed by resides Ala55, Ile96 and Met98 (Fig. 6). The study by Sha and Cao (Sha & Cao, 2015) provided valuable information towards designing novel and potent selective isoform inhibitors of Hsp90. The structures shown in Fig. 6 were taken from the cluster of complex with lower binding energy. The cyclopetane moiety shows different orientation into Hsp90 and Grp94 which also provide basis of isoform selectivity inhibition. Generally, molecular docking has shown to play an important role in discovery of potent Hsp90 inhibitors.

Applications of molecular docking method is limited due to full flexibility of protein, accurate estimation of binding affinity and inclusion of water molecules (few codes are able to include structural water during docking calculations). Such limitations have necessitated the use of all atom molecular dynamics (MD) simulation codes which allow full protein flexibility.

2.5.2 The success of molecular dynamics approaches

Earlier works on molecular dynamics simulations reported by Levitt and co-workers (Levitt & Warshel, 1975) and McCammon and co-workers (McCammon *et al.*, 1977) have provided strong insight into the role of classical MD in simulating biological materials such as proteins and nucleic acids (Levitt & Warshel, 1975; McCammon *et al.*, 1977). In their works, they showed how classical MD can be used to study folding and conformations of proteins and nu-

cleic acids (Levitt & Warshel, 1975; McCammon *et al.*, 1977). Over the recent past years, many researchers have established that MD simulation has a great potential to overcome limitation of static structure based drug design which is mostly applied in docking experiments (Sinko *et al.*, 2013; Kokh *et al.*, 2011; Kim *et al.*, 2018). Protein flexibility is one of the major limitation of common molecular docking tools which (few of them) allow only side chain of the protein to be flexible and this has remained a challenge (Sinko *et al.*, 2013; Kokh *et al.*, 2011; Kim *et al.*, 2018). One of the greatest promise of classical MD in structural based drug design is structural flexibility of protein. Ensemble based docking has emerged as an alternative to static docking (Lin *et al.*, 2002; Amaro *et al.*, 2008a; Lin *et al.*, 2003; Chan *et al.*, 2013; Kim *et al.*, 2018) as detailed in the next subsection.

2.5.3 Relaxed complex scheme

Protein flexibility plays an important role in computational drug design and development. Accounting for receptor flexibility in drug design and discovery has become a subject of interest among medicinal chemists, biophysicists and biochemists. Relaxed complex scheme or ensemble based docking has become an important approach in addressing protein flexibility during docking experiments (Lin *et al.*, 2002; Amaro *et al.*, 2008a; Lin *et al.*, 2003; Chan *et al.*, 2013; Kim *et al.*, 2018). In this approach, a crystal structure of a protein is subjected to a long MD simulation for hundrends of nanoseconds. From the MD structure different snapshots are extracted, clustered and then subjected to virtual screening or molecular docking experiments (Lin *et al.*, 2002; Amaro *et al.*, 2008a; Lin *et al.*, 2003; Chan *et al.*, 2013; Kim *et al.*, 2018). Results from different structures are averaged to obtain improved results reflecting flexible protein structure. This study has utilized RCS aiming at identifying Hsp90 inhibitors using an improved ensemble based docking protocol.

Protein flexibility in drug design and discovery was first reported employing a RCS protocol by the group of McCammon (Lin *et al.*, 2002; Lin *et al.*, 2003). The protocol has gained popularity in structured based drug design (SBDD) where several research groups have used the protocol to account for the protein flexibility in drug discovery regime (Chan *et al.*, 2013; Kim *et al.*, 2018). The protocol has been extended to virtual screening of several compounds (Totrov & Abagyan, 2008). In an effort to discover novel inhibitors of Hsp90, Kim *et al.* (2018) employed an ensemble based docking and biophysical methods to identify potential inhibitors of Hsp90 (Kim *et al.*, 2018). Ensemble based docking identified four compounds (**13-16**) as potential inhibitors of Hsp90. Biophysical studies of the four compounds correlated with ensemble docking results. The two compounds (**13,14**) possessed anticancer activity against MCF7 breast cancer cell, while all four compounds possessed strong inhibitory activities against A549 prostate cancer cell line (Kim *et al.*, 2018). Hsp90 inhibition by the four compounds was checked by measuring the mRNA levels in MCF7 cell. Hsp90 inhibition was

found to be time dependent for all four compounds (Kim *et al.*, 2018). The study by Kim *et al.* (2018) has provided strong bases on developing new and potent Hsp90 inhibitors.

Despite of its success, RCS, still suffers from limitations apart from those of being based on classical MD simulations which are subject to inadequate conformational sampling and force field approximation (De Vivo *et al.*, 2016; Durrant & McCammon, 2011). RCS depends on docking scoring functions which sacrifice accuracy at the expense of achieving high speed (Durrant & McCammon, 2011).

2.5.4 Application of free energies methods

Despite that several docking algorithms are optimized at the expense of time rather than accuracy, however, more accurate methods to predict binding affinity do exist but they are computationally demanding. Methods such as thermodynamic integration, free energy pertubation, single-step pertubation as well as molecular mechanics Poisson Boltzmann (Generalized Born) surface area (MM-PB(GB)SA) are commonly used in calculating binding affinity of ligands to their receptors. In this thesis, TI and MM-PBSA methods by pointing out their success and limitations are discussed.

Free energy calculations using TI are increasingly to variety of biological systems and have aided towards understanding the association and dissociation of protein-ligand complex of interest, in particular Hsp90 with its inhibitors or substrates (Ignjatović *et al.*, 2016; Kawaguchi *et al.*, 2016). Recently, Kawaguchi and co-workers (Kawaguchi *et al.*, 2016), carried out a thermodynamic integration (TI) to investigate the association mechanism of N-terminal domain Hsp90 (ND-Hsp90) and ADP (Kawaguchi *et al.*, 2016). They found that, the association is driven by van der Waals forces of interaction while electrostatic destabilizes the association (Kawaguchi *et al.*, 2016). Furthermore, authors, found that, during the association process, ADP is pulled by Lys69 from bulk region of the pocket and Lys112 at the surface entrance of the pocket and finally pushed out by solvate water molecules in the binding site of ND-Hsp90 (Kawaguchi *et al.*, 2016). Such association process provides a good basis towards understanding the association of protein-ligand complexes as well as towards the discovery of novel ND-Hsp90 inhibitors (Kawaguchi *et al.*, 2016). In another work, the dissociation of ADP from ND-Hsp90 was investigated by using both MD and TI approaches (Kawaguchi *et al.*, 2013). TI revealed the dissociation minimum distance of 0.8 nm with the free energy of -75.2 kJ/mol. Although this binding free energy is contrally to the experimental binding free energy of -26.2 kJ/mol reported by Nilapwar *et al.* (2009) at 20 °C, the difference is due to the inconsistency of the calculated value and the experimental free energy (Kawaguchi *et al.*, 2013). The free energy calculated by Kawaguchi *et al.* (2013) is the free energy difference with the reference state and not standard state. Despite of the difference, the results reported by Kawaguchi *et*

al. (2013), make sense in terms of estimating the binding free energy as a function of distance from reference state. The dissociation of ADP from ND-Hsp90 was observed to be restricted by amino acid Met98 as revealed by MD simulation (Kawaguchi *et al.*, 2013), as the tail group of ADP moves from inside to outside through the side chain of Met98 with a distance of 1.0 nm. Similar observation for ligand movement restriction by Met98 has been previously reported (Immormino *et al.*, 2006).

Flexibility of the ND-Hsp90 binding pocket has remained a challenge in discovering inhibitors not only in experimental approaches but also in theoretical/computational approaches. Several studies have reported the flexibility of ND-Hsp90 binding pocket to be one of the limitations in identifying potential inhibitors (Ignjatović *et al.*, 2016; Simunovic & Voth, 2012; Ribeiro *et al.*, 2018). In their work, Ignjatovic *et al.,* (2016) carried out free energy calculation using alchemical methods, docking and MM/GBSA to estimate the binding affinity of three sets of ligands to Hsp90. The relative free energy was estimated by using the Multi-state Bennett Acceptance-Ratio (MBAR) approach (Ignjatović *et al.*, 2016). Despite the fact that alchemical methods are known to be accurate in estimating the binding affinity (Su & Johnson, 2016; Ytreberg *et al.*, 2006; Aldeghi *et al.*, 2018), free energy results for all sets of ligands were rather disappointing and poorly correlated with the experiment and resulted into larger errors of up 26 kJ/mol (Ignjatović *et al.*, 2016). The reasons for such discrepancy were related to high flexibility of the Hsp90 binding pocket which resulted to; different binding mode in the crystal structure, exclusion and inclusion of structural water and inaccuracy of the molecular mechanics (MM) force field to model the chemical variation of the ligands (Ignjatović *et al.*, 2016). Some of the compounds discovered as Hsp90 inhibitors using MD simulation related methods are presented in Fig. 7.

2.5.5 Application of MD simulation in drug delivery systems

Molecular dynamics and related methods are now widely used in studying the interaction of nanocarriers with their drugs. This subsection points out how molecular dynamics methods have provided insight in studying the interaction of polymers such chitosan nanoparticles and PAMAM dendrimer in drug delivery. The application of different computational tools to study drug delivery systems based on PAMAM dendrimer has been highlighted in our previous work (Shadrack *et al.*, 2018). The current study discusses the application of molecular dynamics in studying drug delivery system based on chitosan nanoparticle.

Polymeric based nanoparticles have shown a great potential in the biomedical field as drug delivery and diagnostic agents. In particular, chitin a 1,4-β linked polymer of N-acetylglucosamine that is cheap and abundantly available in nature (Shan *et al.*, 2014) has gained much attention in biomedical field as drug delivery agent. Chitosan, is a polysaccharide derivative of chitin

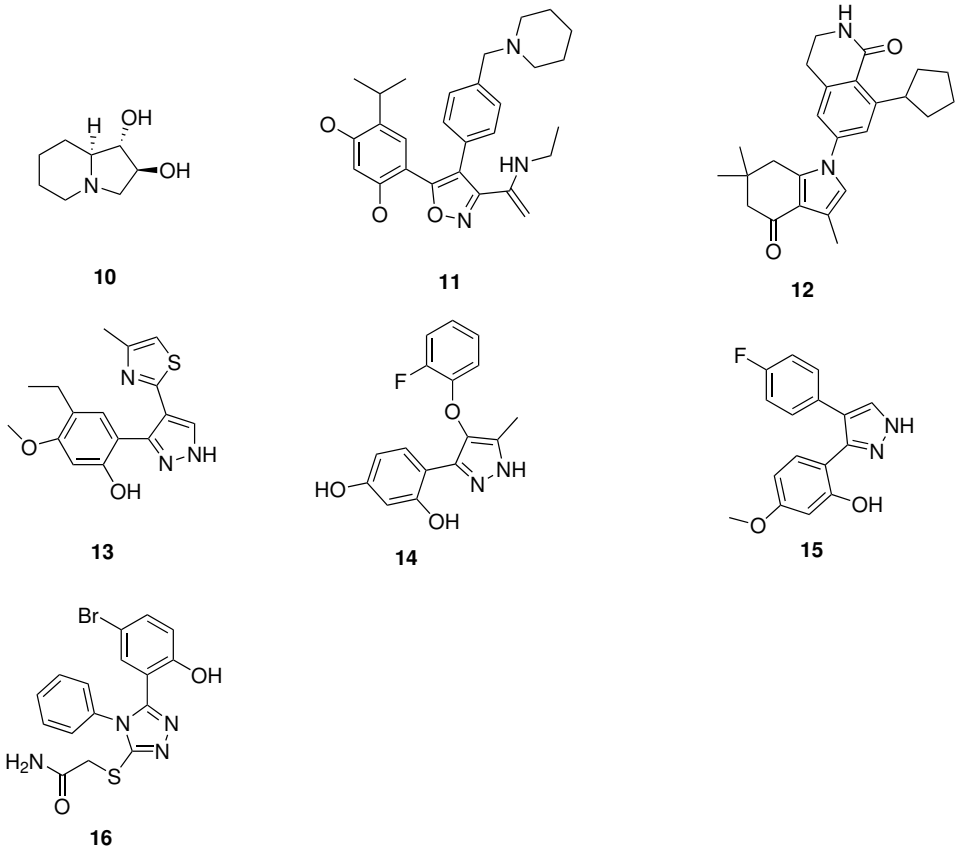


Figure 7: Some of the new discovered Hsp90 inhibitors using molecular dynamics and related methods.

which is obtained by deacetylation process as shown in Fig. 8. Such a process makes chitosan to have improved and desirable properties as a drugs delivery agent than chitin (Shan *et al.*, 2014).

Such improved and excellence properties includes; good biocompatibility, chemical stability, low cost, improved solubility and low toxicity (Shan *et al.*, 2014). The characterization of chitosan-drug interaction for drug delivery system has been widely reported using both experimental (Sanyakamdhorn *et al.*, 2013) and computational (Shan *et al.*, 2014; Shen *et al.*, 2017) methods. Understanding interactions at atomistic level in nanoparticle-drug is of great importance, however, such process is not trivial and rarely gives insights at the atomistic level directly from experiments (Carr *et al.*, 2018). Over the recent years, molecular simulations have aided experimental results by providing fundamental insights into the interaction of host-guest systems. Many successful examples now exist where classical molecular dynamics provided insight at the molecular level to understand the interaction of chitosan and drugs in different solvents (Shen *et al.*, 2017; Razmimanesh *et al.*, 2015; Rungnim *et al.*, 2013). For instance, Shan and coworkers, investigated the interaction of hydrophobic hydrophilic modified chitosan nanoparticles using molecular dynamics simulation (Shan *et al.*, 2014). In their study, MD simulation helped to establish that hydrophobic modified chitosan nanoparticle were suitable for carrying and delivering doxorubicin. In another study, the interaction of chitosan nanoparticles

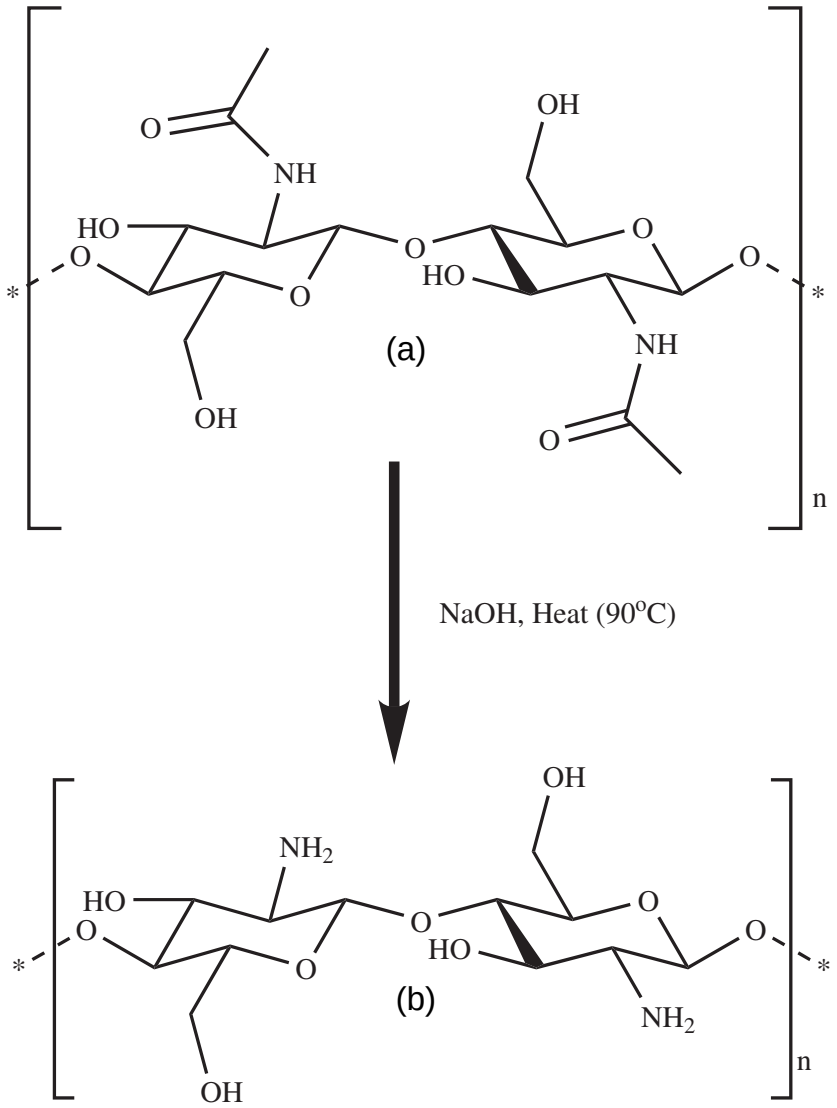


Figure 8: Deacetylation process of chitin. (a) shows the chemical structure of chitin (b) the chemical structure of chitosan.

with polynucleotide at molecular level was investigated using molecular dynamics (Shen *et al.*, 2017). Results presented in their work established and paved the way on how polynucleotides interact with chitosan at different protonation states.

This subsection, highlighted few among many successful and usefulness of molecular dynamics simulation in studying the interactions between chitosan and drug at atomistic level with applications in drug delivery system. Such examples have, however, only based on the thermodynamics aspects with little consideration (if any) on the kinetics and residence time of drug unbinding. This provides a room to further explore the kinetic rate of drug unbinding with its implication to drug delivery systems based on chitosan-drug interaction.

2.5.6 Solution conformation of natural products: Curcumin as a model natural product

Natural products have played a great role in human history and have served as medicine for treatment of many ailments, including malaria, tuberclosis, cancer and inflammation to mention a few (Patsahan *et al.*, 2017; Ilnytskyi *et al.*, 2016). For example, curcumin, isolated from *Curcuma longa* plant, is a well known and documented natural product used as colouring agent, food spices and medicine (Patsahan *et al.*, 2017; Ilnytskyi *et al.*, 2016). Its biological activities range from antiviral, antioxidant, antiiflammatory, anticancer and antidiabetics amongst others (Patsahan *et al.*, 2017; Ilnytskyi *et al.*, 2016; Slabber *et al.*, 2016). Although curcumin possess broad biological activities, its clinical applications are limited due to its poor aqueous solubility (Hani & Shivakumar, 2014). Owing to its numerous biological properties and poor aqueous solubility, curcumin was selected as a model natural products to investigate its solution conformation at molecular level.

It is well documented that, clinical applications of many natural products, and in particular, curcumin are limited due to poor aqueous solubility, chemical instability, low cellular uptake, quick hydrolysis and short half-life (Patsahan *et al.*, 2017; Ilnytskyi *et al.*, 2016; Slabber *et al.*, 2016). In order to address such challenges, a natural product needs to be dissolved in appropriate solvents or need a delivery system which may increase solubility, half-life and cellular uptake (Patsahan *et al.*, 2017; Ilnytskyi *et al.*, 2016; Slabber *et al.*, 2016).

Chemists have well acknowledged that the biological activities of a natural product or small molecule depends on several factors which includes: nontrivial response to solvents, fragment conformation, charge distribution and molecular structure besides other factors (Slabber *et al.*, 2016; Ilnytskyi *et al.*, 2016). Investigation of structural conformation helps in understanding the pharmacological properties of a natural product in question. Despite the fact that, curcumin has been studied for many years, its solution conformation has given little attention with few studies been reported using computational and experimental methods (Hazra *et al.*, 2014; Patsahan *et al.*, 2017; Ilnytskyi *et al.*, 2016; Slabber *et al.*, 2016).

Because of some limitation to experimental methods in investigating materials at molecular level, computational simulation approaches have provided a better way of understanding behaviour of materials at molecular level. In order to understand the behaviour i.e configuration, conformation, stability and orientation preferences of curcumin in different solvents, in this thesis, computational methods in particular metadynamics an enhanced sampling method was used to explore such properties and behaviour. In addition, as discussed in Chapter 1, Section 1.1.4 of this thesis, the role of solvents on residence time and kinetics of host-guest complex are investigated using computation methods to provide the underlying information on how solvents affects the interaction. Two solvents are used *viz.* DMSO and water, DMSO is a clinical

relevance solvent and some drugs formulated with DMSO exist (Slabber *et al.*, 2016).

2.6 Conclusion and way forward

This Chapter has discussed different computational methods employed in drug discovery and design. The chapter started by discussing various computational methods from molecular docking, molecular dynamics, metadynamics and free energy methods (TI and MMPBSA). Some successful examples where such methods have been applied in the field of drug design have been highlighted. Although different docking protocols and molecular dynamics are used in many studies, there is a lack of study which describes the effects of inclusion of water into different docking protocols such as docking to crystal structures with and without water, docking to ensemble holo and apo protein structures with and without water. On the other hand, the use of molecular dynamics to study the interaction of chitosan-drug interaction is well appreciated, however, there is lack of studies (if any) which report the effect of different solvents on the kinetics and residence time of drug unbinding. The effects of different solvents on conformation and kinetics of natural products is also an area in drug design which needs to be given attention. Studies of heterogeneous interaction at crystal-water interface using molecular dynamics simulation in drug design has gained less attention. Many studies have focused on the thermodynamics properties on drug binding, with less attention payed to kinetics, residence time, interaction at the crystal structure-water interface and polymophism changes which have a great implication to drug design.

In this study different protein-ligand docking protocols are explored. The thesis further reports the effect of solvents on the conformation and kinetics of small molecules using curcumin as a model molecule. The solvent effects on kinetics and residence time on host-guest unbinding is also investigated using toussantine-A-chitosan complex as a model system. The next chapter describes the material and methods used in this thesis.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Molecular docking

3.1.1 Protein preparation

The X-ray crystallographic structure of Hsp90β (PDB ID: 3NMQ) (Yun *et al.*, 2011) in complex with EC-44 inhibitor was obtained from RCSB Protein Data Bank. For validation purpose, the complexes of Hsp90α (PDB ID: 4EGK) (Austin *et al.*, 2012), Grp94 (PDB ID: 3O2F) (Patel *et al.*, 2013) and Grp94 (PDB ID: 1QY8) (Soldano *et al.*, 2003) were also obtained from RCSB Protein Data Bank (PDB). For molecular docking calculations the proteins and ligands were prepared using FlexX (part of LeadIT) molecular docking software, BioSolveIT, GmbH, German (Kramer *et al.*, 1999). The binding site of the protein was examined from the crystal structures using EC-44 as a reference ligand (Kramer *et al.*, 1999) and finally confirmed by X-ray data from the literature (Stebbins *et al.*, 1997). After identifying the binding site, the ligand was removed from the active site and the protein was subjected for virtual screening.

3.1.2 Validation of molecular docking protocol

Before inferring any prediction from the docking procedure it is important to understand its limitations. For example, one of the limitation of docking methods to predict experimental results, is the accuracy of the scoring function (David *et al.*, 2005). One of the best known approach to validate the docking program is its ability to retrieve the binding mode of known drugs in the protein. In this study, root mean square deviation (RMSD) and experimental binding energies were used to validate the docking scoring function. RMSD is used as the measure to indicate if the binding mode is successful or not. If the RMSD is less or equal to 2.5 Å, then, the scoring function is able to reproduce and predict the binding mode for other drugs (Brooijmans, 2009; Hevener *et al.*, 2009). Ligands were extracted from their binding modes and redocked again and their RMSD was assessed and presented in Table 1 in Chapter 4. This validation confirms the goodness of the results reported in this thesis. In particular, the RMSD of the redocked ligands was $\leqslant$ 2.5 Å. In a similar way, experimental data from literature (Baum *et al.*, 2009; Baum *et al.*, 2010) obtained by using isothermal titration calorimetry method and the calculated binding free energy reported in this thesis were compared and found to be in a reasonable agreement with the correlation value r^2 = 0.92 as shown in Figure 10 Chapter 4.

3.1.3 Virtual screening

Small molecules (drugs) used in this study were obtained from drug bank database, which comprise of more than seven thousands approved small molecules and macromolecules (Wishart

et al., 2007). Small molecules were downloaded with their chemical structure in SMILES (Simplified Molecular Input Line-Entry System). The SMILES were then converted into .sdf and .mol file format. The Lipinski (Lipinski *et al.*, 1997) rule of five (LRo5) was used in pre-screening (filtering) the small molecules. A total of 2000 small molecules were obtained and subjected to virtual screening against Hsp90β. Virtual screening was done by an in-house environment in Mmcule with the Autodock Vina (Trott & Olson, 2010) docking program. Pyrrolopyrimidine methoxypyridine (EC-44) a known Hsp90 inhibitor with a binding energy of -9.2 kcal/mol was used as a reference ligand. 26 approved drugs with binding energies less than that were chosen and subjected to relaxed complex scheme molecular docking (Fig. 9).

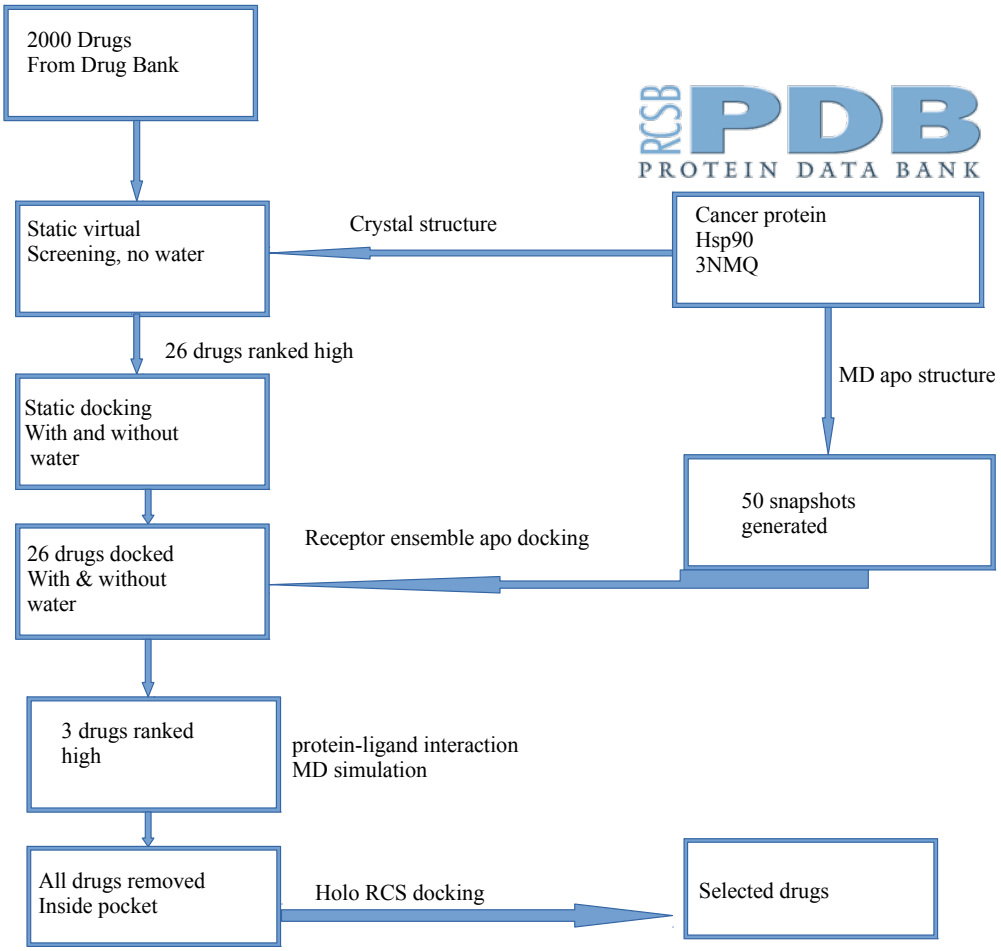


Figure 9: Drug repurposing protocol employing the relaxed complex scheme used in this study.

3.1.4 Molecular docking with FlexX

FlexX (LeadIT) is a flexible docking method which employs an incremental construction (IC) algorithm to dock ligands into the active site (Kramer *et al.*, 1999). In FlexX, the ligand is decomposed into many components, then, protein interactions guide the reconstruction (assemble of ligands) into the active site by using various placement strategies (Kramer *et al.*, 1999). Ligands were docked into the active site. Docking calculations were performed on different types

of conformations: crystal structure with and without water and receptor ensembles for apo and holo structures from the MD simulation with and without water. For the crystal docking protocol, water was chosen using a cutoff of 6.5 or 10 Å from the center of mass of the reference ligand EC44.

The binding free energy was calculated and assessed by using HYDE (HYdrogen bond and DEhydration energy in protein-ligand complexes) scoring function (Schneider *et al.*, 2013) as described in Chapter 2. During docking, a total of ten poses were generated and assessed for their binding energies. Poses with lower energy for each docked conformation were recorded and averaged as mean relaxed complex (RC) binding energies.

3.2 Molecular dynamics (MD) simulations

3.2.1 Protein-ligand interactions

Molecular dynamics simulations were carried out with GROMACS (GROningen MAchine for Chemical Simulations) Ver. 5.1.2 (Abraham *et al.*, 2015) with the amber03 force field (Wang *et al.*, 2004). The crystal structure of the Hsp90 (PDB ID: 3MNQ) (Yun *et al.*, 2011) with ligand removed, was used to generate apo structures for the relaxed complex scheme. Protein topologies were generated using GROMACS while ligands topologies were generated using antechember (Wang *et al.*, 2004). A total of four proteins system simulations were performed corresponding to apo and holo (bound) systems. The systems were solvated with the TIP4P (Horn *et al.*, 2004) water model and sodium ions were added to neutralize the system. All the systems were energy minimized using steepest descent algorithm with position restraints. Energy minimized systems were then equilibrated for 5 ns then followed by a production run for 250 ns and 100 ns for apo and holo systems, respectively. At the equilibration stage, the temperature and pressure was maintained using the Berendesen method (Berendsen *et al.*, 1984). For the production stage, Parinello-Rahman (Parrinello & Rahman, 1981) was employed for pressure coupling at 1 bar and v-rescale (Bussi *et al.*, 2007) was used for temperature coupling at 300 K. Particle Mesh Eward (PME) (Cheatham *et al.*, 1995) was used to treat the long-range electrostatic interactions. Covalent bonds were constrained by using LINCS algorithm (Hess *et al.*, 1997) and a time step of 2 fs was used for all calculations. For all systems (apo and holo structures), after the protein reached equilibrium, configurations were sampled every 2 ns for the docking experiments. A total of 50 configurations were used for the docking.

3.2.2 Curcumin in bulk solvents

The crystal structure of curcumin was extracted from the Cambridge Crystallographic Data Center. A single molecule of curcumin was differently solvated with pre-equilibrated solvents *viz*: 1620 TIP4P (Jorgensen *et al.*, 1983) water molecules, 682 methanol (MeOH), 572

dichloromethane (DCM), 505 carbon tetrachloride (CCl_4) and 402 dimethyl sulfoxide molecules, to have five different systems. The solvated systems were energy minimized using steepest descent algorithm to remove any overlap of the atoms from the staring coordinates. All systems were then equilibrated at NPT ensemble as previously reported for a single curcumin in water (Ilnytskyi *et al.*, 2016) at a temperature of 300 K and 1 bar using a time step 0.1 fs for 500 ps. A production run was carried at NPT ensemble with a temperature of 300 K and 1 bar for 5 ns using Parrinello-Rahman barostat (Parrinello & Rahman, 1980). Bonds were constrained using LINCS (Hess *et al.*, 1997), while Particle Mesh Ewald (PME) (Darden *et al.*, 1993) methods were used for handling long-range electrostatic interactions. All simulations were performed using GROMACS code (Abraham *et al.*, 2015) version 2016 with OPLS-AA (Jorgensen *et al.*, 1996) force field. The configuration from MD simulation was used for metadynamics simulation.

3.2.3 Chitosan-TouA supramolecule in solvents

For host-guest interactions, the initial structure of chitosan was obtained from PubChem database and functionalized with triphosphate (TPP) groups. Group functionalization and energy optimization were done using Avogadro (Hanwell *et al.*, 2012) employing the MMFF94 force field. The structure was then subjected to molecular docking to obtain the chitosan-TouA complex. Docking calculations were performed using ArgusDock docking engine (Oda *et al.*, 2007) to obtain the complex with the lowest binding energy for MD simulation.

MD simulation was done using GROMACS 2016 (Abraham *et al.*, 2015) employing GROMOS 54a7 (Schmid *et al.*, 2011) force field. The topologies were generated using PRODRG (Schüttelkopf & Van Aalten, 2004). All systems were solvated in explicit solvents, DMSO and SPC (Gereben & Pusztai, 2011) water model . All the systems were energy minimized using steepest descent algorithm with position restraints. The systems were equilibrated at NVT for 500 ps and NPT ensemble for 1 ns, respectively, then followed by a production run for 140 ns. During the equilibration stage, both temperature and pressure were maintained using Berendsen method (Nosé, 1986). For production stage, Parinello-Rahman (Parrinello & Rahman, 1981) and v-rescale (Bussi *et al.*, 2007) were used for pressure and temperature coupling at 300 K, respectively. Particle Mesh Ewald (PME) (Cheatham *et al.*, 1995) was used to treat long-range electrostatic interactions while covalent bonds were constrained using the LINCS (Hess *et al.*, 1997) algorithm, a time step of 2 fs was used for all calculations. The MD configuration was then used for metadynamics simulation to investigate the kinetics and residence time.

3.3 Well-Tempered metadynamics (WT-MetaD)

Solvent-solute interaction, drug unbinding kinetics and residence time were investigated using well-tempered metadynamics simulation (Barducci *et al.*, 2008) as described in Chapter 2.

During WT-MetaD the biasfactor and Gaussian height were set to 10 and 1.2, respectively, the sigma values were obtained from short unbiased MD simulation. The following CVs were chosen to describe the conformational changes of curcumin in vacuum and bulk solvent, CV1: the keto-enol distances, CV2: end to end distance, CV3: torsional angles and CV4: number of hydrogen bonding. While, for chitosan-TouA complex interaction, the following CVs were used, CV5: minimum distance between chitosan and TouA selected from specific groups for both chitosan and TouA, CV6: coordination number and CV7: torsional angle were used to describe the unbinding process in different solvents. All WT-MetaD and free energy analyses were done using plumed plugin for MD simulation version 2.4 (Bonomi *et al.*, 2009).

3.3.1 Residence time (τ_{AB}) and k_{off} calculations

In this thesis, drug residence time from state A to B (τ_{AB}) is defined as the time taken by TouA to reach the solvent-exposed surface of chitosan from the bound state. A $d_{min} > 0.6$ nm was used as a criterion for TouA to have completely unbound and solvent-exposed as at this distance there is no interaction between chitosan and TouA. The time exchange from state A to B and the transition state/activation barrier are denoted as τ_{ex} and TS, respectively. Then, τ_{AB} for the unbinding process that is from bound state A to unbound state B in the two solvents is computed as:

$$\tau_{AB} = \tau_{ex} \cdot \exp^{(\mathrm{B}(TS-FA))} \tag{3.1}$$

where B is the Boltzmann constant which is 2.6 kJ/mol and *FA* is the free energy for bound state.

3.4 Binding free energy calculated by MM-PBSA

The MM-PBSA calculations were applied for three complexes of Hsp90-pitavastatin, Hsp90-ezetimibe and Hsp90-vilazodone obtained from the MD simulation runs. The binding free energy was performed using *g_mmpbsa* tool implemented in GROMACS (Kumari *et al.*, 2014). A total of three hundred equally spaced snapshots were obtained from a single MD run, water and counterions were removed before the calculations started. In this implementation, the ΔE_{MM} calculated in gas phase is the total of van der Waal and electrostatic interaction energies. The $\Delta G_{\mathrm{solvation}}$ is calculated as the sum of polar and non-polar. The polar part (ΔG_{PB}) was calculated by linearly solving the Poission-Boltzmann equation, with a grid spacing set to 0.5 Å. The exterior dielectric constant (ε) was set to 80 and different values for solute dielectric constant (ε) =1,2,3,4 were used. The non-polar (ΔG_{SASA}) was determined using the solvent accessible surface area (SASA) with a gamma value set to 0.0226 and calculated as 0.0226 x $\Delta SASA$. In *g_mmpbsa* all the energy terms are calculated separately and then added up to obtain the contribution of each residue to the binding energy.

3.5 Thermodynamic integration

In order to further accounts for the effects of solvents on chitosan-TouA stability, binding free energy based on thermodynamics integration (TI) was performed. TI is one of the most effective equilibrium approach for calculating the binding free energy (Kirkwood, 1935; Ytreberg *et al.*, 2006). In TI the change in free energy is performed at different value of λ, as follows:

$$\Delta\Delta G_{\text{binding}} = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\delta H}{\delta \lambda} \right\rangle_{NPT,\lambda} d\lambda \tag{3.2}$$

where $\langle ....\rangle d\lambda$ represents an ensemble average at different values of λ. TI was performed on 21 equally spaced λ value of 0.0, 0.05, ...0.95, and 1.0. The first 500 ps were discarded for equilibration, then Equation. 3.2 was used to estimate the data using a trapezoidal rule. The convergence of the system is well monitored and described in appendix.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 The role of water and conformational fluctuations in Hsp90 in response to inhibitors

4.1.1 Molecular docking benchmark and validation

Before embarking into molecular docking simulations, one has to validate the method used if it can produce reliable results. In this thesis, RMSD as well as comparison to experimental data were used to validate the docking method as described in Chapter 3. The RMSD of the redocked ligands and their binding energies as compared to experimental data are presented in Table 1 and Fig. 10. Both RMSD and correlated binding energies showed that the method is reliable in reproducing docking poses and can be trusted for other docking experiments.

Table 1: RMSD (Å) of the X-ray crystal structure and redocked ligands of the best conformations in their active sites.

PDB Code	Complexes	FlexX RMSD (Å)
3NMQ	Hsp90β-EC44	0.78
4EGK	Hsp90α-Radicicol	0.29
3O2F	Grp94-P54-1	1.99
1QY8	Grp94-Radicicol	0.25
2FWZ	Hsp90α-PU-H71	2.55
2EXL	Grp94-Geldanamycin	1.80
2GFD	Grp94-Radamide	1.81

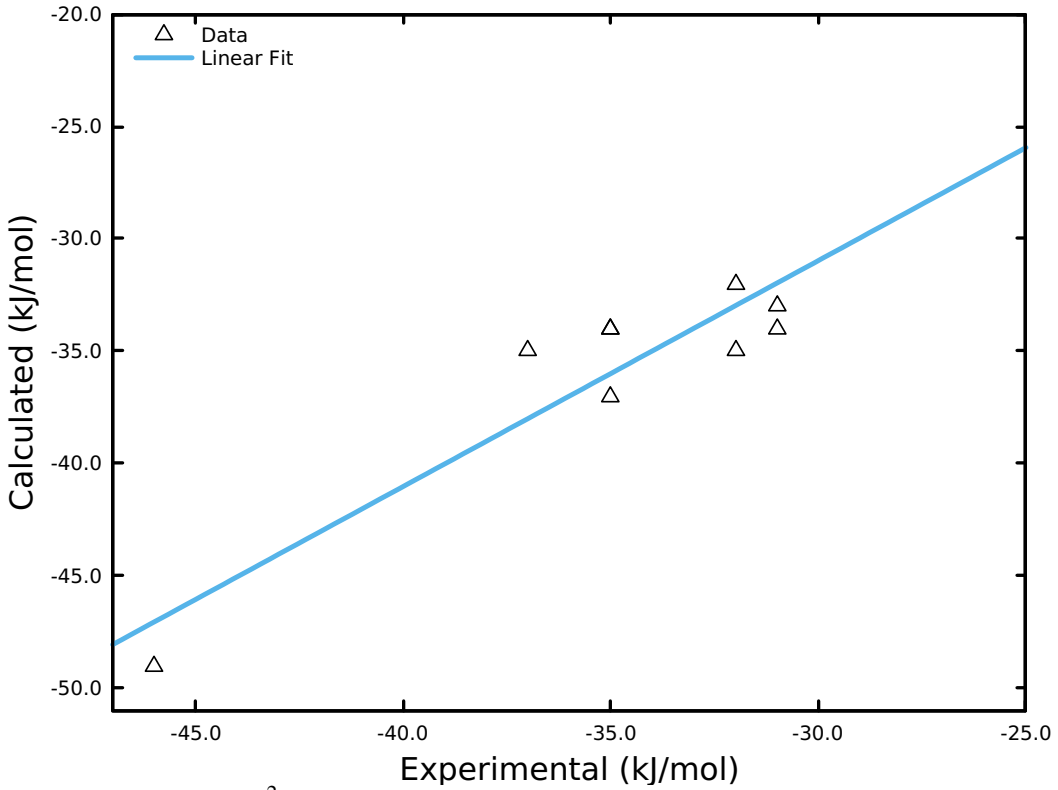


Figure 10: Correlation ($r^2 = 0.92$) of experimental and calculated binding free energy (kJ/mol).

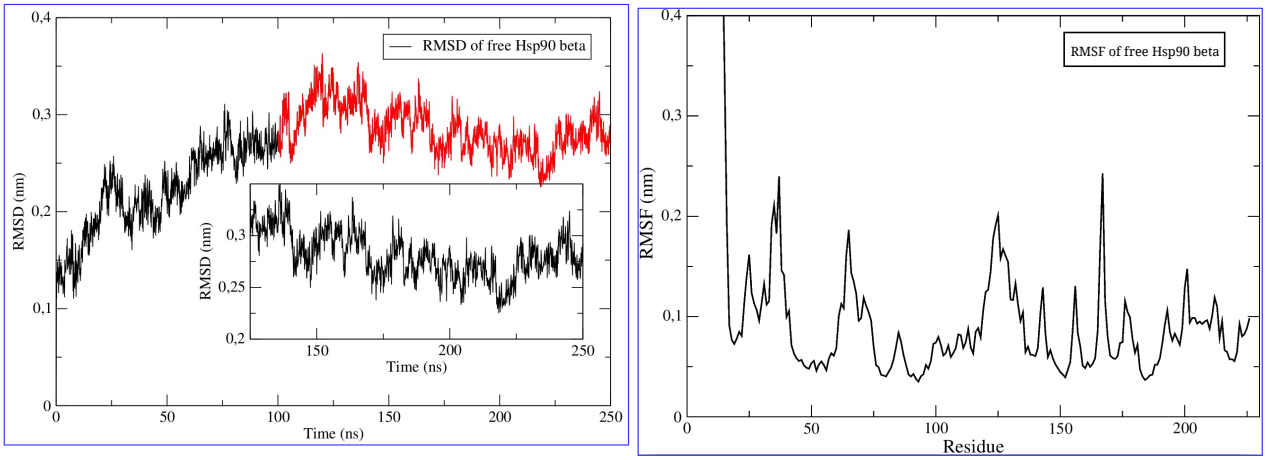
4.1.2 Molecular dynamics and conformation changes of Hsp90β active site

Accommodating full protein flexibility in docking calculations is challenging. Many docking programs do not allow full protein flexibility during molecular docking calculations. To adopt full protein flexibility, the RCS approach was employed. 250 ns run molecular dynamics simulation of Hsp90β was carried out to reveal the most flexible parts of the protein, to generate flexible receptor ensembles for molecular docking and to understand flexibility/movement of residues and water molecules in the active site. It was hypothesized that pocket fluctuation would result in different binding modes of the drugs and changes in numbers of water molecules. Some drugs showed different binding modes and energies when docked to different MD snapshots. Some snapshots were observed to have large ΔG, while some snapshots had small ΔG which resulted in a large variation of the binding free energies. The effects of protein pocket fluctuations on binding free energies (ΔG) of drugs have also been reported in some studies which involved docking to MD snapshots (Hernández-Rodríguez *et al.*, 2016).

The backbone RMSD of the free (unbound state) Hsp90β (Fig. 11a) equilibrated after 100 ns and remained stable until the end of the run with an RMSD mean value of 0.26 nm. Further analysis showed large fluctuations corresponding to helix 4 (α4) with amino acids from 105 to 114 as shown in Fig. 12. Fluctuations on helix 4 have also reported in the literature (Barker *et al.*, 2010; Prodromou *et al.*, 1997). Hsp90 has been reported to adopt conformation change at ATP pocket on the N-terminal domain, and it was reported that the nature of such conformation remained unknown (Barker *et al.*, 2010; Krukenberg *et al.*, 2011; Prodromou *et al.*, 1997). High fluctuations of helix 4 also affect the binding pocket, and this has remained one of the challenges in designing Hsp90 inhibitors (Barker *et al.*, 2010; Krukenberg *et al.*, 2011; Prodromou *et al.*, 1997). The flexibility of this region is also reported for Hsp90α (Yan *et al.*, 2008). Besides RMSD the root mean square fluctuation (RMSF) of the protein was calculated. The RMSF indicated more oscillation/fluctuations in the region with amino acid residues 60-75, 120-130 and 165. Such fluctuations were less than 0.2 nm (Fig. 11b). Other regions involved in the binding sites remained stable.

The N-terminal domain (NTD) of the Hsp90 contains 226 amino acid residues made up of α helixes and β strand sheets. The α helixes (H2, H4, H7) and Loop1 constitute the ATP binding pocket and make up the wall of the ATP pocket, while the bottom is made up of the antiparallel β strand sheets (β3, β4, and β7). The entrance diameter of the ATP binding pocket has a diameter of 12 Å, depth 15 Å and 8 Å midway diameter as it narrows down (Kawaguchi *et al.*, 2016; Stebbins *et al.*, 1997). The entrance surface of the pocket is made up of polar solvent accessible amino acid residues Lys58, Lys69, Asn106 and Lys112. As it narrows down the pocket it is formed by the hydrophobic, charged and polar amino acid residues: Leu48, Asn51, Asp54, Ala55, Ile91, Asp93, Ile96, Gly97, Met98, Asn106, Leu107, Gly135, Phe138, Val150,

Thr184 and Val186 (Stebbins *et al.*, 1997).



(a) (b)

Figure 11: (a) General RMSD of free Hsp90β over 250 ns, the inset shows sampling time from 100-250 ns, (b) the RMSF of free Hsp90β taken from the equilibrated MD run

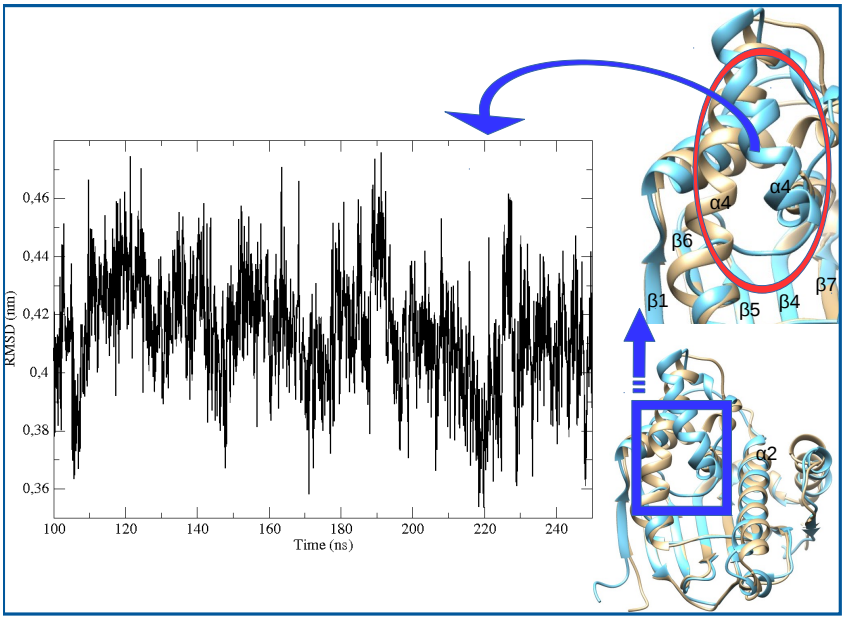
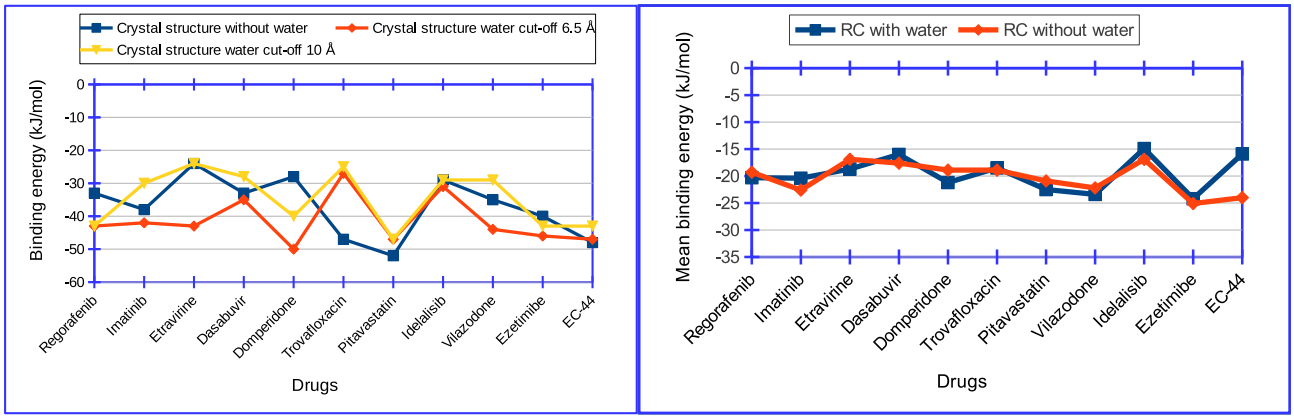


Figure 12: Fluctuation of alpha helix 4 with amino acids from 105-114. Light blue shows the movement of alpha helix 4 from its original position at the beginning of the simulation.

4.1.3 Relaxed complex scheme (RCS)

To gain more understanding on the role of protein flexibility and water in molecular docking, a relaxed complex scheme (RCS) approach defined in Fig. 9 was used. All 26 drugs obtained from virtual screening were docked to protein conformations (receptor ensembles) extracted from the molecular dynamics simulation. Figure 13 shows the docking results of drugs docked to the X-ray crystal structure and MD snapshots with and without water for some of the drugs.



(a) (b)

Figure 13: (a) Sensitivity of drugs when docked to crystal structure without water (dark blue), with water cut-off 6.5 Å (red) and water cut-off 10 Å (yellow). (b) Sensitivity of drugs to RC structures with and without water.

Docking calculations of all 26 drugs to X-ray crystal structures with and without water showed different binding energies (Fig. 13a). Some drugs showed lower binding energies when docked to crystal structure with water as compared to without water. For example, trovafloxacin showed a difference of -23 kJ/mol for a crystal structure with and without water. The large difference in the binding energy could be explained by the nature of the drug as it contain more hydrophobic groups which would find it unfavourable to have water molecules in the pocket during molecular docking and hence less binding energy for crystal structure with water. Similar reasons hold for drugs such as imatinib and idelalisib which the inclusion of water molecules make the binding energy less favourable and displayed a large difference in binding energy. Some drugs did not show significant different in energy when docked to protein with and without water.

The sensitivity of results to the crystal structure by including more water by changing the cut-off tha is 6.5 and 10 Å was also investigated. The criteria for choosing a 6.5 Å was from the reference molecule EC44, a co-crystallized inhibitor of Hsp90. The cut-off of 10 Å was decided to understand if the inclusion of more water will have effects on the binding free energies. Docking to the crystal structure with water using a criteria cut-off of 6.5 and 10 Å showed similar binding energy except for vilazodone, etravirine and domperidone (Fig. 13a). Using vilazodone as an example, it can be said that the inclusion of several water has effects on the binding energy of drugs and different drugs require different number of water for optimal functioning (Santos *et al.*, 2009). When vilazodone was docked using criteria of 6.5 Å lower binding energy of -44 kJ/mol was obtained, however using a 10 Å cut-off criteria the binding energy reduced to -29 kJ/mol.

When docking was performed using receptor ensembles with and without water small differ-

ences in binding energy were recorded as shown in Fig. 13b. In some few cases water showed an effect to some drugs, for instance, docking of dicoumarol showed lower binding energy when docked to receptor ensemble with water compared to receptor ensembles without water. Other drugs showed that water-mediated their interaction with protein while in other drugs, water did not cause changes in binding energy. Some drugs, for instance, trovafloxacin, behaves in a peculiar manner, when it was docked to receptor ensembles with and without water did not show a difference in binding energy, however, when docked to crystal structures large differences in energy is observed. In few cases, drugs showed a similar trend for docking to receptor ensembles and crystal structures with and without water. In their study, Santos et al. showed that the effect of water on mediating protein-ligand interaction depends on the nature of drug and protein conformation (Santos *et al.*, 2009). While some studies (de Graaf *et al.*, 2006) have shown that water does not improve the docking energy of drugs, in this study it has been found that, the effect of water depends on the nature of drugs and the conformation of the protein. In general, it is observed that there is sensitivity of docking results when using crystal and ensemble structures whereas, RCS seems to reduce the sensitivity of the docking results with or without water.

4.1.4 Protein-ligand complex MD simulation

In order to gain more understanding on the stability of the docked drugs and the protein-ligand complex, a molecular dynamics simulation was performed on the three drugs; ezetimibe, vilazodone and pitavastatin to understand their dynamics inside the Hsp90 pocket over the simulation time. The three ligands selected were least sensitive to the docking protocols. MD simulations for each protein-drug was monitored by checking the energetics and structural properties of each complex for the whole simulation period of 100 ns. The structural stability of the complexes was examined by the root mean square fluctuations (RMSF) and the distances between the protein and the drugs. The following subsections present and discuss the residue fluctuations induced by the ligand upon binding as well as the minimum distances between the protein and drugs.

4.1.5 RMSF analysis

In order to capture detailed information on residue fluctuations induced by the ligand upon binding, the RMSF difference between holo and apo protein is computed and presented in Fig. 14. One can see that all the drugs induced conformational changes to the residues but to different extent. For example, in all structures one can observe residue fluctuations at the region with residues from 20-40, 120-140 and 170. Residue fluctuations for vilazodone and pitavastatin at the region from 20-40 exhibit similar nature with little changes compared to ezetimibe. On closer look, ezetimibe and vilazodone shows fluctuations at the region from 60-70, which are not observed in pitavastatin. The region made up of 120-140 and 170 residues shows similar

fluctuations for pitavastatin and vilazodone which are not observed in ezetimibe.

In order to provide a predictive picture on the origin of these effects for the three drugs an illustration of protein-ligand interaction complexes is presented in Fig. 15. It is observed that the region from 20-40 residues is composed of the N-terminal loop and beta sheet 1 (β1), high fluctuations at the beginning are expected due to high fluctuation of the loop. It is further observed that, all drugs interact with alpha helix (α1) region made up of residues 22-36. Vilazodone and pitavastatin show similar fluctuations which are not observed to ezetimibe, this is because both vilazodone and pitavastatin are seen to bound near these residues and enhance fluctuations as indicated in Fig. 15. However, ezetimibe is hanging up between loop (L4) and loop 5 (L5) and is not able to interact with regions formed by residue 120-140, instead it interacts with alpha helix 3 (α3).

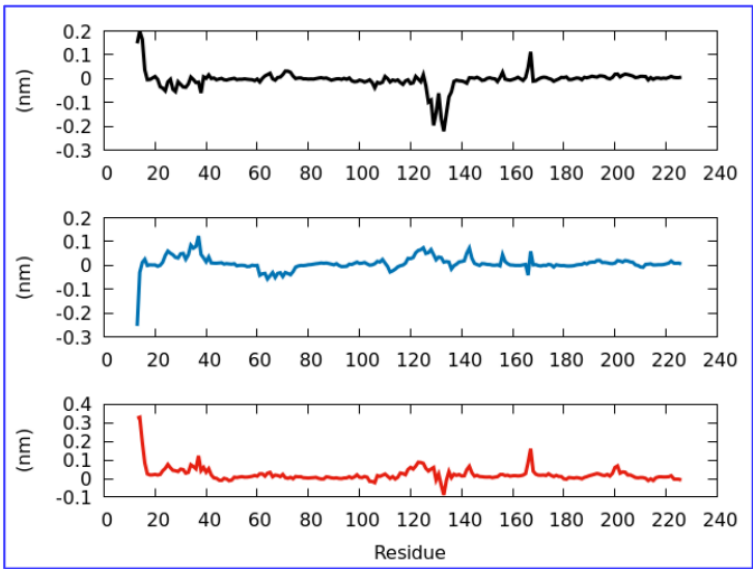
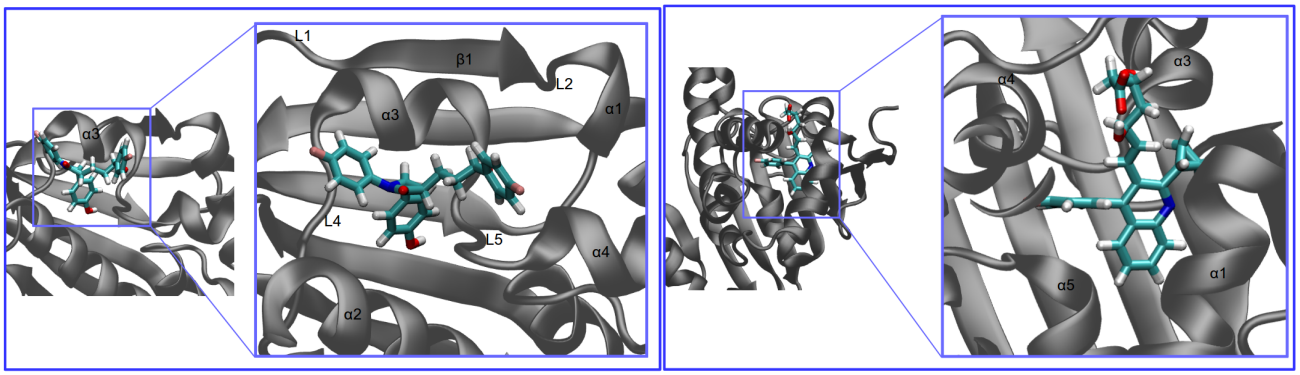


Figure 14: RMSF differences of holo and apo protein. Vilazodone is shown in black colour, ezetimibe in blue and pitavastatin in red colour.



(a) (b)

Figure 15: Binding orientations of the two drugs inside the Hsp90 pocket. (a) is the binding mode for ezetimibe and (b) is the binding mode for pitavastatin.

4.1.6 Minimum distances between ligands and the protein

In order to provide quantitative information of the movement of the drugs inside the protein pocket, the minimum distance between the ligands and the protein was measured. Figure 16 shows the minimum distances of each groups of the ligand during the simulation time. For clarity, group 1 is coloured black, group 2 is coloured red and group 3 is colored in blue as indicated in the left pannel of Fig. 16. A general observation for all ligands is that, pitavastatin shows the smallest minimum distance of less than 0.25 nm for all groups compared to vilazodone and ezetimibe. The reason for this is that, pitavastatin is buried inside the pocket and it interacts with the residues formed by alpha helix 1 (α1), 4 (α4) and (α5). On the other hand, the two groups of vilazodone shows similar distance with group 3 being larger. A larger distance of group 2 and 3 is attributed by the fact that these two groups are more exposed to the solvent and are freely to move while group 1 is buried inside and its movement is restricted. Group 3 of ezetimibe also displays large distances compared to the other groups as it points to the entrance of the pocket and is exposed to the solvent. However, the two groups point towards the beta sheets (group 2) and helix 4 (group 1) where their movement is restricted.

4.1.7 Holo RCS cross docking

The previous subsections discussed how the three drugs change protein conformation and the fluctuations of the drugs in the protein. In this subsection, the cross docking experiment on holo protein structures obtained from the apo docking experiment is discussed. Cross docking refers to the process of docking a ligand to a protein containing different molecule (Sutherland *et al.*, 2007). In holo cross docking, all the drug candidates are removed from their bound state and then each ligand is docked to three different protein ensembles from holo-vilazodone, holo-pitavastatin and holo-ezetimibe. From cross docking experiment it is observed that, holo structure of pitavastatin possessed lower binding energy close to crystal structure (Fig. 17). Lower binding energy of the holo structure of pitavastatin is explained by its binding mode inside the protein. As depicted earlier in Fig. 16, minimum distance analysis shows that pitavastatin has a smallest distance for all the groups suggesting that it is buried inside the protein pocket and enhances little fluctuation to the protein. However, this is not observed for vilazodone and ezetimibe, for example, groups 1 and 2 of vilazodone and group 1 of ezetimibe show larger distance suggesting that they are not buried and hence resulting to large fluctuation of the protein. Such large fluctuations of the protein reduce the ability of the ligand to correctly bind into the active site hence less binding energy is obtained.

4.1.8 Binding free energy calculated by MM-PBSA

Binding free energies calculated by MM-PBSA are effective in complimenting results obtained from docking and MD simulation experiments. In this thesis, the binding free energy for the three complexes was calculated using MM-PBSA method. Although the method is effective in

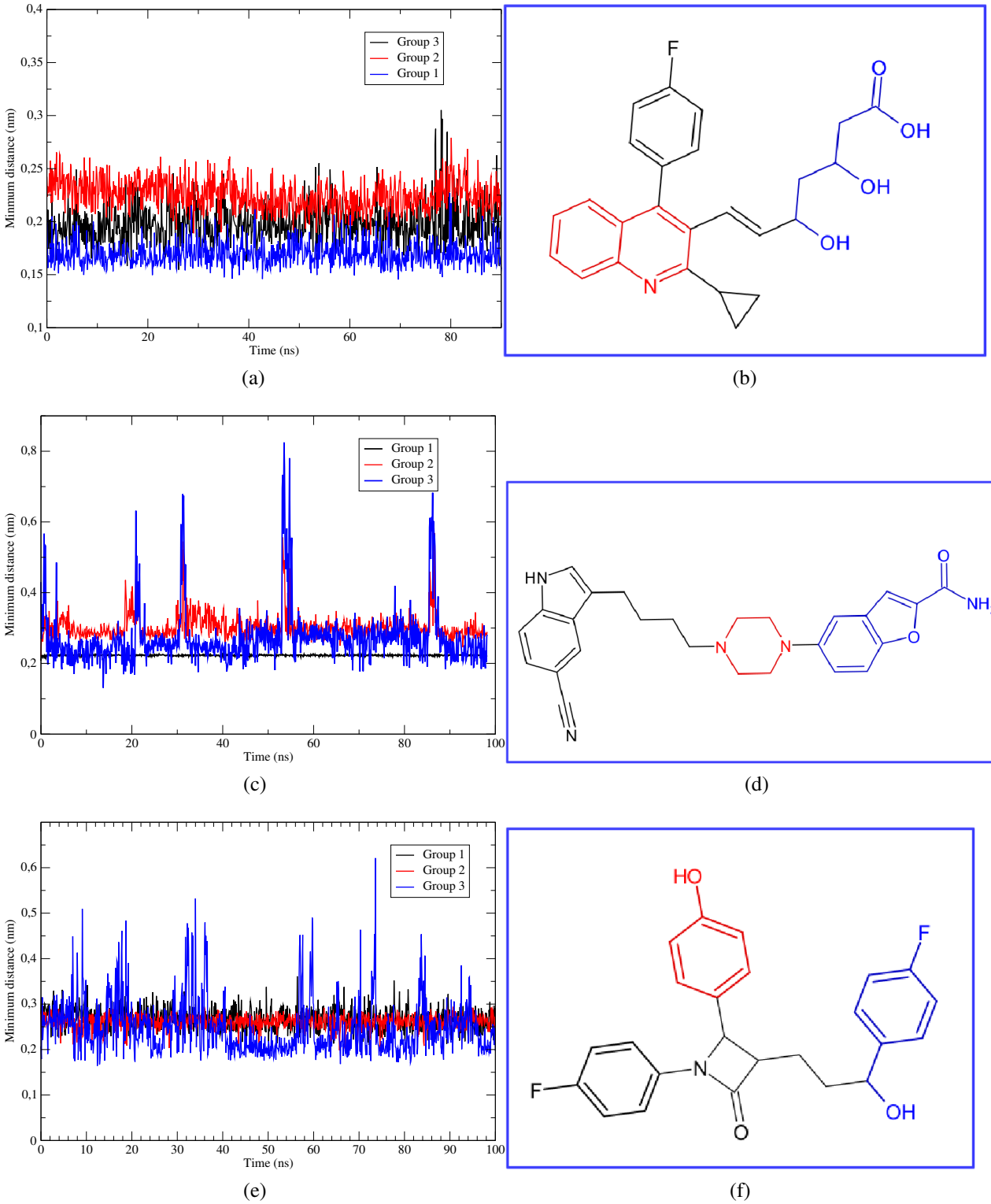
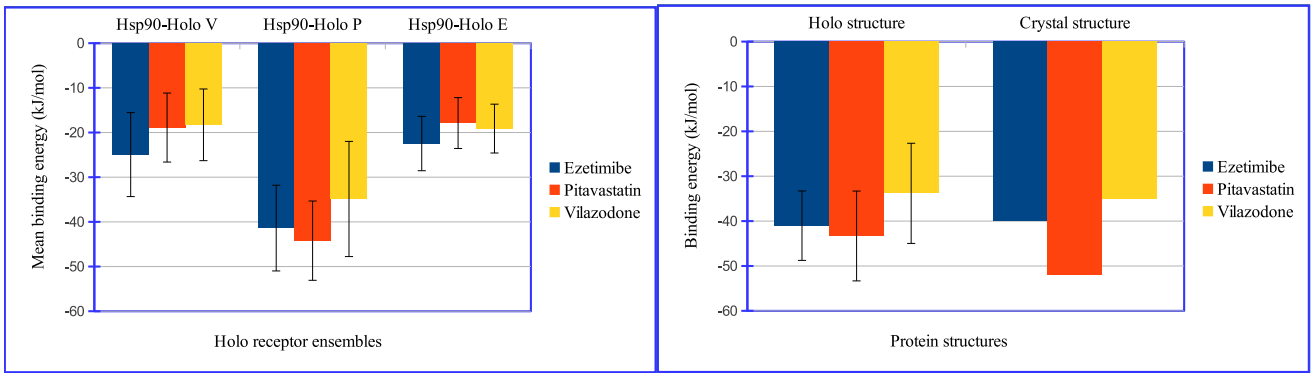


Figure 16: Minimum distances between protein and the three ligands with their respective groups. Colours shown on the minimum distance (left panel) represents the groups for each ligand shown in the right panel. For all ligands group 1 is shown in black, group 2 in red and group 3 in blue.

estimating binding free energy in protein-ligand interactions, accuracy of the results obtained by this method depends on solute dielectric constant (ε), nature of the protein active site and



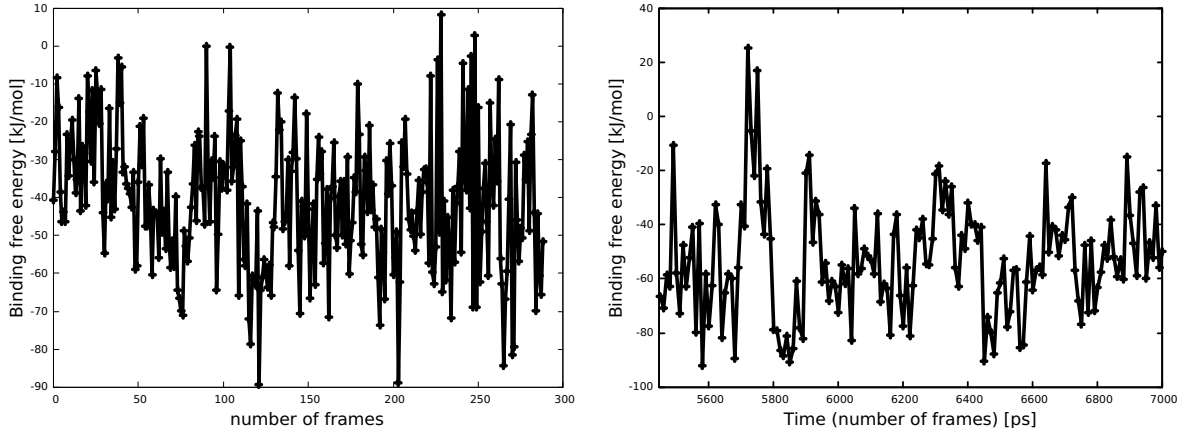
(a) (b)

Figure 17: Binding energy of 3 drugs in different holo structures (a) Docking into holo structures of vilazodone (V), pitavastatin (P) and ezetimibe (E) without water (b) Comparison in binding energy between the crystal structure and holo structure of pitavastatin.

length of the simulation time, usually, a smaller time is convenient. In this thesis, the binding free energies were calculated from a single trajectory. The convergence of the free energies for the complexes is presented in Fig. 18, as has shown in previous reports (Hou *et al.*, 2010) that the short MD run usually yields good results when compared to long MD run, in this work a short MD run also yielded good results compared to long simulation. Furthermore, different values for solute dielectric constant (ε) ranging from 1-4 were tested and the $\varepsilon = 4$ yielded reasonable result. This is because of the nature of the binding pocket which contains hydrophobic and charged residues as it narrows down and such properties require large value of ε. Table 2 shows the binding free energies obtained by MM-PBSA. The results suggest a strong binding affinity for pitavastatin, followed by ezetimibe. However, a weaker interaction for vilazodone complex was observed. The MM-PBSA results further support the RCS and MD simulation results where pitavastatin complex shows to be more thermodynamic stable. Electrostatic and van der Waal forces play an important role is stabilization of the complexes, in pitavastatin electrostatic interaction highly contributes to stabilization of the complex. This is in agreement with the MD results where pitavastatin is burried and remain stable in the Hsp90 pocket. In addition, non-polar energies also contribute to stabilization of the Hsp90-pitavastatin complex when compared to other complexes (Table 2). The observed larger error values is due to the intrinsic properties of MM-PBSA method which result into large errors. Larger error values in MM-PBSA have also been observed in previous reports (Peräkylä & Nordman, 2001). In addition, large errors could also be attributed by large conformation changes at the ATP binding pocket as previously observed for the docking results. The binding free energies obtained from MM-PBSA further suggests that pitavastatin is the most stable and can bind strongly to inhibit Hsp90 activities.

The energy term contributions listed in Table 2 show that van der Waals (ΔE_{vdW}), electrostatic (ΔE_{ele}) and non-polar ($\Delta E_{non\text{-}polar}$) favorably contribute to the binding free energy, while polar

energy terms (ΔE_{polar}) are repulsive and unfavorably contribute to the binding free energy for all ligands. Comparing the energy term contributions, one can see that the ΔE_{vdW} energies are larger and correlate with the size of the ligand i.e pitavastatin with topological polar surface area of 90.6 $Å^2$ is larger than ezetimibe whose topological polar surface area is 60.8 $Å^2$. Large size of hydrophobic molecule has been also observed to largely contributes to ΔE_{vdW} energy terms (Weis *et al.*, 2006; Kuhn & Kollman, 2000). The $\Delta E_{non\text{-}polar}$ energy terms are observed to be generally small but with correlation to the size of the ligand. The large value for ΔE_{vdW} energy is ascribed to the nature of the protein active site and the chemical nature of the ligands. The Hsp90 active site contains charged and hydrophobic residues as it narrows down the active site, the ligands also have more hydrophobic groups that are interacting with the protein. As an example, Fig. 16a-b, shows that the more hydrophobic groups in pitavastatin are more buried and interacting with the hydrophobic residue of the protein pocket. Thus, the interaction between Hsp90 and ligands is more dominated with hydrophobic ΔE_{vdW} interactions. On the other hand, it can be said that polar groups of the ligand did not find adequate bonding terms in the protein active site. Large ΔE_{vdW} and repulsive or unfavorable ΔE_{polar} contribution to the binding energy in Hsp90 has also been previously reported for more hydrophobic molecules (Sha & Cao, 2015). Other studies have also reported large domination of the ΔE_{vdW} energy terms and the reason for such domination was the nature of the small molecules where the charged molecules were observed to have large domination of the ΔE_{vdW} energy terms and resulted to more repulsive polar energy terms for biotin molecules binding to avidin protein (Weis *et al.*, 2006; Kuhn & Kollman, 2000).



(a) (b)

Figure 18: Convergence of binding free energies at different snapshots for ezetimibe and pitavastatin-Hsp90 complexes calculated by MM-PBSA method, (a) ezetimibe and (b) pitavastatin.

Table 2: Relative binding free energies (kJ/mol) from MM-PBSA decomposition.

System	ΔE_{vdW}	ΔE_{ele}	ΔE_{polar}	$\Delta E_{non\text{-}polar}$	$\Delta E_{binding}$
Hsp90-ezetimibe	-123.3	-45.3	142.4	-14.8	-41.1 ± 17
Hsp90-pitavastatin	-222.2	-48.9	142.6	-22.5	-53.2 ± 20

4.2 Solvent effects on drug interaction and conformation

This section presents and discuss the effects of different solvents on the properties of drugs/small molecules in bulk solvents. In particular the section focuses on discussing the properties and conformation of curcumin in bulk solvents and solvent effects on drug kinetics and residence time using chitosan-TouA complex as a model system. The thesis begins by discussing the properties and conformation of curcumin in bulk solvents.

4.2.1 Properties and conformation of curcumin in bulk solvents

Curcumin (Fig. 19), is a natural product derived from rhizome roots of the *Curcuma longa* plant. Curcumin exists in keto-enol tautomerism. In its crystallographic form, the enol tautomer is the most stable (Ilnytskyi *et al.*, 2016; Patsahan *et al.*, 2017; Slabber *et al.*, 2016) and is of most interest in this study. Since drug formulation takes place in organic solvents, understanding solution conformation of curcumin is an important step towards designing curcumin formulation for various biological applications. In this thesis, metadynamics, an enhanced sampling method is used to study the solution conformation of curcumin in different solvents.

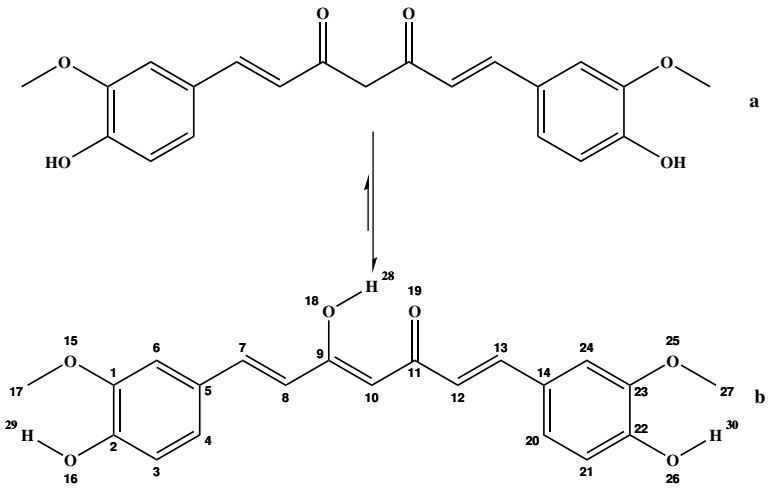


Figure 19: Chemical structure of curcumin molecule. Keto form of curcumin is indicated in top (a) and the enolic form is below (b).

4.2.2 Curcumin in vacuum

To account for the effects of solvent in curcumin conformation and orientation, the behaviour of curcumin in a vacuum is investigated by biasing keto-enol distance, ring-ring distance and spacer dihedral angle connecting the two rings. The thesis begins by describing the free energy profile of the keto-enol distance and orientation preferences in a vacuum. Figure 20 (a) shows the 1D plot for keto-enol distance with three energy minima at 0.19, 0.37 and 0.5 nm. The free energy minima at 0.19 nm correspond to the enolic hydrogen pointing towards the keto oxygen. The free energy minimum at 0.37 and 0.5 nm corresponds to the enolic hydrogen of

the *cis* and *trans* conformation pointing towards and away from the keto oxygen, respectively. The tendency of enolic hydrogen to point towards keto oxygen is expected due to strong intramolecular electrostatic interactions between the enol hydrogen and keto oxygen.

The ring-ring distance is obtained by measuring the center of mass distances of the two rings. The ring-ring distance shows flat free energy with a minimum at 0.6 to 0.8 nm (Fig. 20b). A deep minimum with free energy at a distance of 1.17 nm is observed which suggests that, although there is a bent conformation, curcumin exist in a stretched conformation in a vacuum. The observed distances and conformation changes are due to dihedral angle changes at the spacer region connecting the two rings formed by atoms 8, 9, 10, 11 and 12, 11, 10, 9 and are named as A_1. The second dihedral angle group near the rings is formed by atoms 10, 9, 8, 7 and 10, 11, 12, 13 these dihedral angel are named as A_2.

To gain further understanding of the origin of these conformation changes of curcumin in a vacuum, dihedral angle distribution at the spacer region are analysed and presented in Fig. 20c and d. The dihedral angle at the spacer region A_1 shows a minimum at $\phi = 0$ and $\phi = \pm 3$ rad, indicating the presence of both *cis* and *trans* conformations. However, the dihedral angle at A_2 shows the minimum at $\phi = \pm 3$ rad showing the presence of *trans*-conformation only. The distribution of dihedral angles suggests the presence of three conformations *viz trans-trans*, *trans-cis* and *cis-cis* conformations. The next subsection discusses the effect of different solvents on curcumin conformation and orientation preferences.

4.2.3 Curcumin in solvents

This subsection discusses the conformation and orientation preferences of a single curcumin molecule in bulk solvents *viz.* water, methanol (MeOH), dichloromethane (DCM), dimethyl sulfoxide (DMSO) and carbon tetrachloride (CCl_4). First, the discussion begins by exploring the 1D free energy profile for curcumin keto-enol distance in different solvents. The free energies are computed from the converged time dependence system for both keto-enol distances and end to end distances. As an example, appendix 36 shows the diffusivity and convergence of the systems for end to end distances in the studied solvents. The discussion begins by exploring the free energy obtained by biasing the keto-enol distance. Figure 21 and Table 3 show the free energy profiles and the difference in free energy (ΔF) values for keto-enol distances in solvents, respectively. As shown in Fig. 21, the effect of solvents is manifested by different minimum depth for the keto-enol (H28-O19) distances. DMSO a solvent where curcumin dissolves shows a small difference in ΔF value followed by water, CCl_4, MeOH and DCM which shows a large ΔF value (Table 3). The small ΔF value in DMSO and water explains the tendency of enolic hydrogen to point outward the keto oxgyen group. As an example, in water due to its electronegativity tendency, water would prefer to form intermolecular hydrogen bond

with the enolic groups causing a C-O bond rotation which results for the enolic group to prefer pointing outward the keto group.

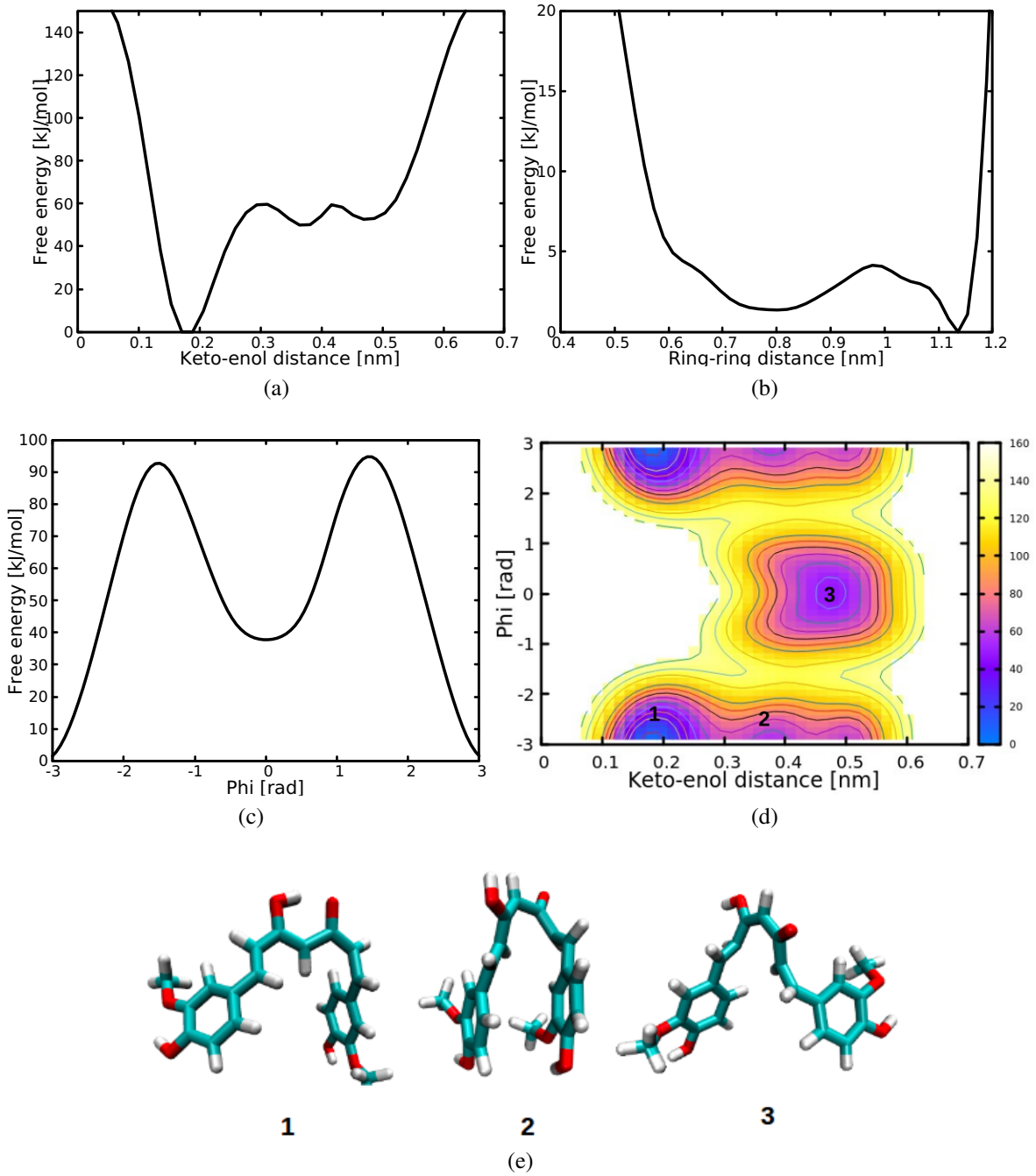


Figure 20: a-c, 1D conformation of curcumin in vacuum. (a) keto enol distance, (b) ring-ring distance, (c) Dihedral angle distribution of the spacer connecting the two rings of curumin in vacuum. d-e, 2D free energy landscape with representative snapshots for *cis-trans* conformation (d = 0.19 nm), *cis-cis* conformation (d = 0.37 nm) and *trans-trans* conformation (d = 0.5 nm) labeled as 1, 2 and 3, respectively.

This is contrary to less polar solvents and in vacuum where a large ΔF value is observed as a result of intramolecular hydrogen bonding interaction between the keto-enol groups. On the other hand, larger ΔF value indicates the tendency of enolic hydrogen to point towards the keto group.

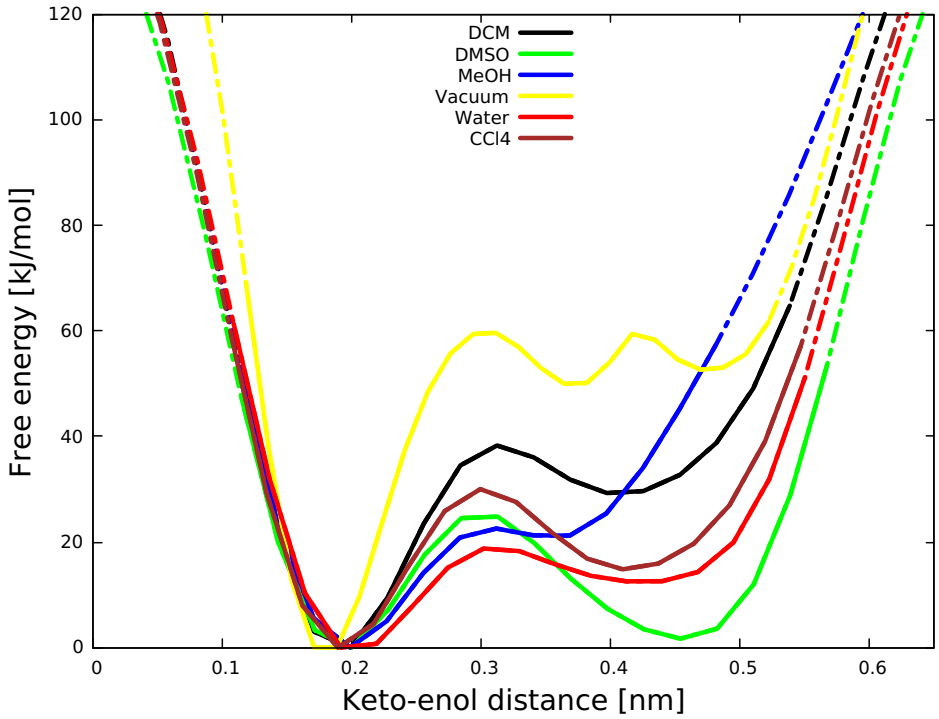


Figure 21: 1D free energy for keto-enol distances in different solvents. The effect of solvent is manifested by different minima.

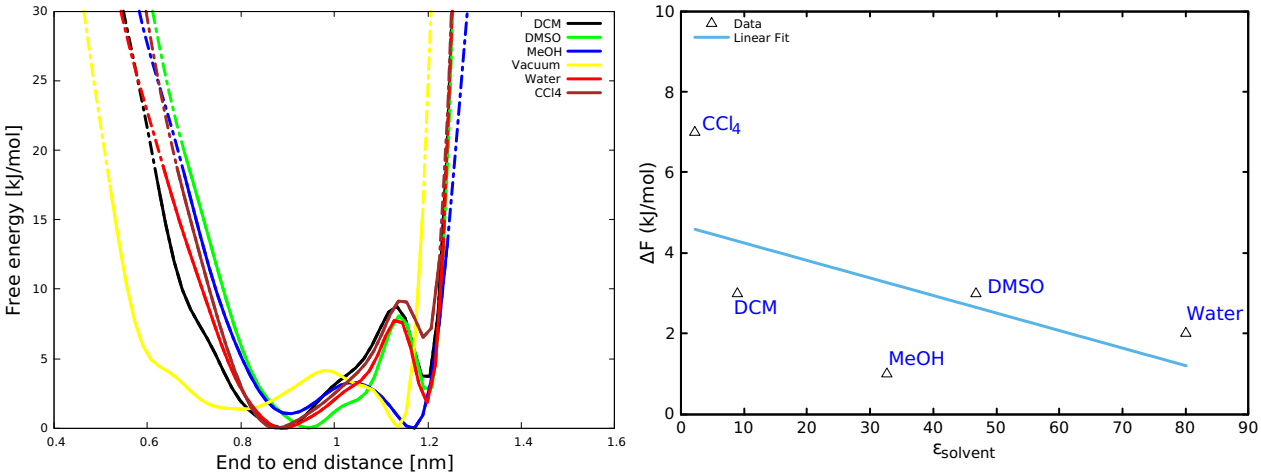
Table 3: Free energy differences (ΔF) (kJ/mol) for keto-enol distances in different solvents.

Solvents	ΔF
DMSO	1
Water	12
CCl_4	15
MeOH	22
DCM	29
Vacuum	53

These results suggest that curcumin with enolic hydrogen pointing towards keto oxygen is the most thermodynamic stable in organic solvents capable of solubilizing it. Although there is lack of experimental data on the solution conformation of curcumin except that of Slabber and co-workers, the present result provides basic information on understanding the solution behavior of curcumin in different organic solvents. Previous works on curcumin model studied by standard molecular dynamics, for example, Pizio *et al.* (2017) using OPLS-UA force field observed two maxima of curcumin in methanol at 0.19 and 0.37 nm for the keto-enol distance, while Samanta and Roccatano (2013) using GROMOS96 force field did not observe the second maxima at $\approx$ 0.4 nm in methanol.

To understand the effects of solvent on the end to end distance, the end to end distance was biased by defining as the distance between the centers of mass of carbon composing the rings

and results are presented Fig. 22 and Table 4. The free energy profiles for all solvents show minima at 0.9 and 1.17 nm but with varying depth for bent and stretched conformation. To understand the influence of solvent on curcumin conformation, the free energy difference (ΔF) for all solvents presented in Table 4 shows that polar solvents such as water, methanol and DMSO have smaller ΔF value compared to non-polar solvent such as CCl_4 which has larger ΔF value. DMSO and MeOH are organic solvents capable of solubilizing curcumin, smaller ΔF value in these solvents is expected as the free energy changes between bent and stretched conformation is small, however, in water where curcumin has marginal solubility the small ΔF is due to the hydrophobicity property, as it will tend to avoid the contact with water and tends to bent. To describe the influence of solvents on conformational changes of curcumin, liquid dielectric constant is correlated with the ΔF value (Fig. 22b) which indicates that more polar solvents with higher dielectric constant such as water have small change in free energy as compared to less polar and /or non-polar such as CCl_4 whose dielectric constant is lower. These results further suggest that liquid dielectric constant influences conformation changes of curcumin in a peculiar manner. Previously, using standard molecular dynamics simulation with OPLS-UA force field and SPC/E water model for single curcumin molecule, Ilnytskyi *et al.* (2016) reported the maxima conformation for ring-ring distance of $\approx$ 1 nm. The presented biased results show conformation with the minima for ring-ring distance at 0.9 and 1.2 nm with varying depth.



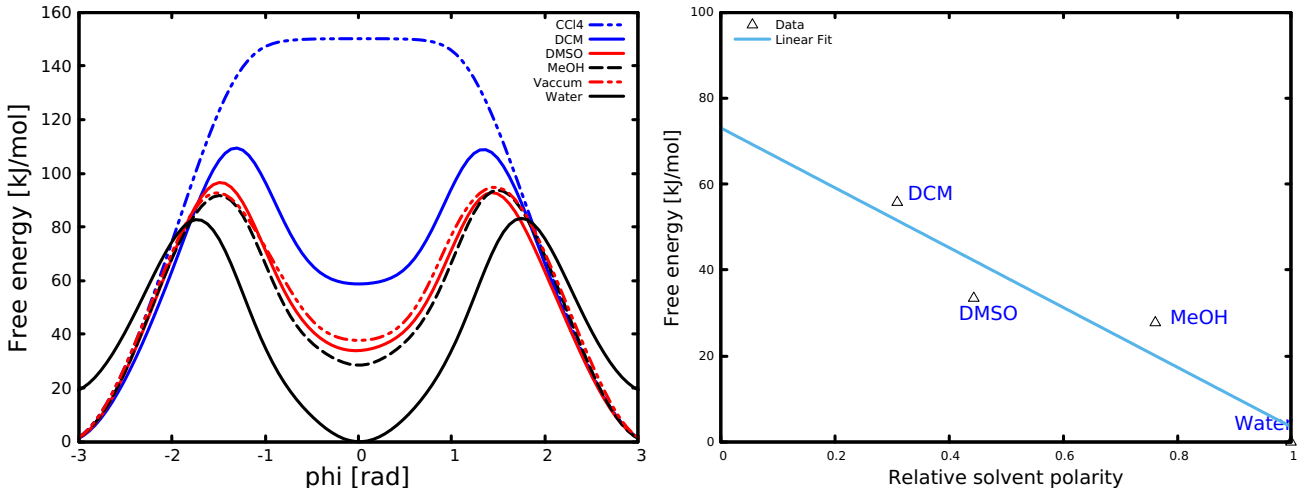
(a) (b)

Figure 22: (a) 1D free energy for end to end distances showing the minimum distances in different solvents. (b) Shows the correlation ($r^2 = -0.59$) between liquids dielectric constant (ε) and free energy differences (ΔF) on curcumin conformation changes.

Table 4: Free energy differences (kJ/mol) for curcumin end to end distances in different solvents.

Solvents	ε_{solv}	ΔF
MeOH	32.70	1
Water	80.10	2
DMSO	46.70	3
DCM	8.93	3
CCl_4	2.24	7
Vacuum	-	2

To understand the origin of these conformation changes in different solvents, dihedral angles at the spacer region define as A_1 and A_2 (see Fig. 19) were biased. The measured dihedral angles at A_1 show the spacer connectivity between the two ring groups and provide information regarding curcumin rotation. Figure 23a shows the 1D free energy profile for dihedral angles of curcumin in different solvents. One can see a clear effect of solvent on the dihedral angle distribution, two peculiar behaviours for curcumin in the solvent under study are observed, first CCl_4 is observed to have a ϕ rotation at $\pm$ 3 rad only, while in other solvents a rotation at ϕ = 0 and $\pm$ 3 rad is observed, the effect of solvents is observed with different free energy minima. The rotation at $\phi = \pm$ 3 rad indicates the presence of *trans*-conformation and at ϕ = 0 indicates *cis*-conformation. In another way, in CCl_4 only the *trans*-conformation is observed while in other solvents both *cis* and *trans*-conformations are present. The *cis*-conformation in water is more thermodynamic stable followed by MeOH, DMSO and least in DCM. It is observed that solvent polarity play an important role in conformation changes and stability. The *cis*-conformation is more stable in polar solvent and the stability decreases as the polarity decreases as shown in Fig. 23b.

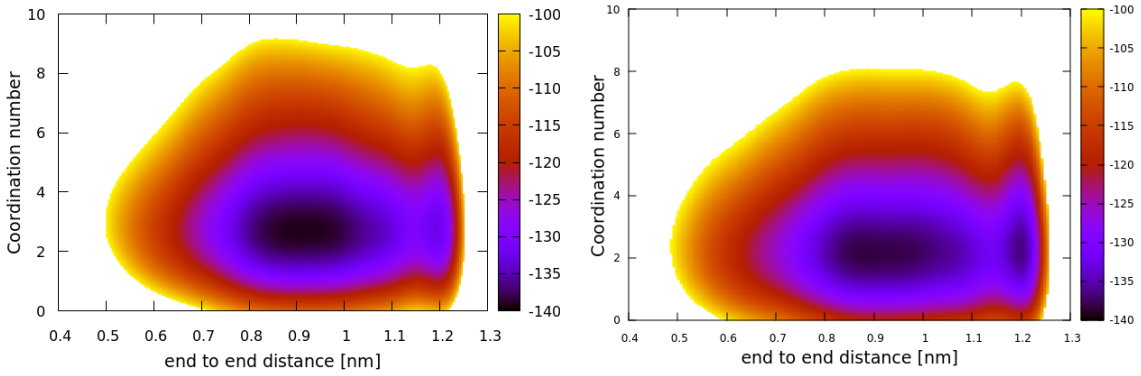


(a) (b)

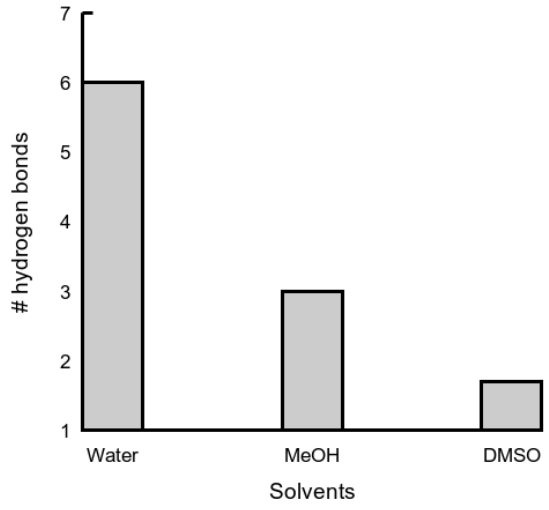
Figure 23: (a) 1D free energy on dihedral angle distributions. (b) The correlation ($r^2 = -0.95$) between relative solvent polarity and free energy for *cis*-conformation. CCl_4 is removed for clarity of the plot as it has larger free energy of 149 kJ/mol.

4.2.4 Hydrogen bonds

Solute-solvent interaction involving donor or acceptor group results into formation of hydrogen bond and can portray kinetic information of the solute molecule. In order to understand the effects of solvents on the hydrogen bond formed by curcumin and vicinal solvents, hydrogen bond was computed by biasing the donor and acceptor groups in solvents and curcumin. Oxygen groups in water and MeOH both can act as donor or acceptor while that in DMSO only acts as donor. The contribution of acceptor or donor groups to hydrogen bonding in polar solvent is nearly the same. As an example, in water the number of hydrogen bonds formed when water acts as an acceptor or donor with the oxygen or hydrogen of curcumin molecules is almost the same (Fig. 24b). However, the total number of hydrogen bonds formed (Fig. 24c) is different in each solvent, with water having larger number of hydrogen bonds and least in DMSO.



(a) (b)



(c)

Figure 24: Hydrogen bonds formed between curcumin and water calculated using donor/acceptor groups. (a) water is donor (b) water is acceptor. (c) total number of hydrogen bonds formed in water, MeOH and DMSO.

It is observed that solvent polarity plays an important role on hydrogen bond formation, as the polarity decreases the number of hydrogen bonds also decreases in the order of water $>$ MeOH $>$ DMSO. The number of acceptor groups play an important role on the formation of hydrogen bonds. Water and MeOH are observed to have higher number of hydrogen bonds due to the presence of an acceptor and donor groups, however, in DMSO this is not the case as the oxygen in DMSO only acts as a donor group and can only form a hydrogen bond with acceptor groups of curcumin molecule. The three solvents portray different hydrogen bonding properties due to the difference in their polarity and presence of acceptor and donor groups.

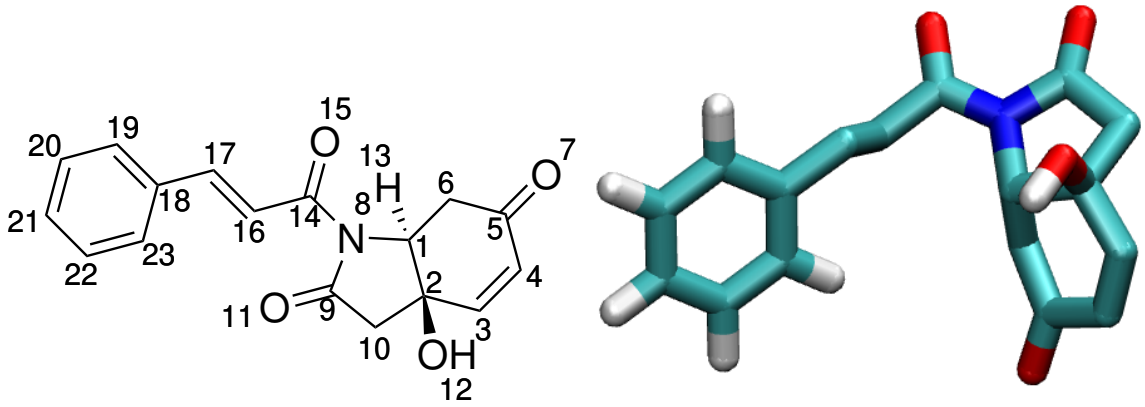
4.3 Solvent effects on host-guest kinetics and residence time

Many complex biomolecules processes such as protein folding, aggregation, drug (un)binding kinetics amongst others, usually take places in aqueous solutions and solvents play an important role in aiding such processes (Jong *et al.*, 2017). The effect of solvent in aiding such processes usually occurs in a non-trivial manner. As an example, supramolecular drug interactions usually take place in solvents which in-turn affect their interaction and eventually their kinetics and residence time. Understanding such processes at molecular level provides a better way in establishing the related solvent effects for such processes. The motivation of this thesis was further to investigate the related solvents effect on drug kinetics and residence time at an atomistic level based on chitosan nanoparticle and toussantine-A.

Toussantine-A (TouA) shown in Fig. 25 is a natural product with promising biological activities including antitubercular and anticancer (Nyandoro *et al.*, 2015). TouA was first isolated and reported by Nyandoro *et al.* (2015), despite its promising biological activities, the molecule showed chemical instability in various solvents during isolation. One of the proposed ways to address the challenges of chemical instability is the use of supramolecular host-guest interaction (Shan *et al.*, 2014). In this approach, a drug is loaded into a host which is a nanoparticle and transported or released to the specific site of action. Thus, polymers such as chitosan due to their good biocompatibility, chemical stability, low cost, improved solubility and low toxicity (Shan *et al.*, 2014) have shown great potential in drug delivery systems. In this thesis, due to the relevance that, drug formulation, synthesis and encapsulation, usually take place in solvents (Carr *et al.*, 2018), the solvent effects on the kinetics and residence time on the unbinding process using TouA and chitosan nanoparticle as a model system are studied.

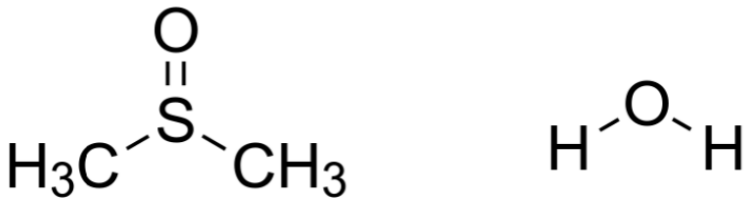
4.3.1 Structural changes of chitosan nanoparticles in DMSO and water

This subsection of the thesis begins by exploring the structural changes of chitosan nanoparticle brought by the solvents. The solvents used *viz.* DMSO and water are polar aprotic and polar protic, respectively, and have clinical relevance (Slabber *et al.*, 2016). Their chemical structures are presented in Fig. 25c. The structural changes of chitosan nanoparticle in the



(a) TouA 2D

(b) TouA 3D



DMSO

Water

(c)

Figure 25: The chemical structures (a) 2D and (b) 3D representation of TouA, atoms are labeled by color, oxygen is red, nitrogen is blue, hydrogen is while and carbon in light blue. (c) 2D representation of DMSO and Water.

two solvents is investigated by measuring the solvent accessible surface area (SASA) which provides information on the influence of solvent on structural changes. SASA provides information on the segregation and aggregation of polymers and protein as well as describing the number of unfavourable hydrophobic contacts between the protein and or polymers in solvents. SASA values are measured from MD simulation of 100 ns and their results are presented in Fig. 26. The SASA value for DMSO is higher when compared to that in water which suggests the formation of segregated structures, while lower value suggests the presence of aggregated structure in water. The aggregation of chitosan nanoparticle in water is explained by the presence of hydrophobic groups which tend to aggregate when are in contact with water (Rampino *et al.*, 2013; Philippova *et al.*, 2001; Wang *et al.*, 2013; Franca *et al.*, 2011). The size distribution of chitosan nanoparticle in the two solvents is further described by calculating the free energy of the SASA values as $F = -k_{\mathrm{B}}T\ln(P)$, where P is the probability and k_{B} is the Boltzmann constant and T is temperature. As shown in Fig. 26 the size of chitosan nanoparticle in DMSO is 24.2 Å larger than in water which is 16.5 Å. The SASA results in the two solvents show that chitosan nanoparticle has a larger size in DMSO than in water which may facilitate more DMSO to enter into the nanoparticle and weaken the interaction hence TouA in DMSO could be released more easily. To further ascertain this the number of solvent near TouA bound to chitosan was measured using a distance cutoff of 3.5 Å. It was observed that few water molecules

were present inside the chitosan which stabilize the interaction by formation of hydrogen bond (Fig. 27), however, for DMSO system it was observed that more DMSO molecules enter into the nanoparticle and destabilize the interaction as DMSO lacks an acidic hydrogen.

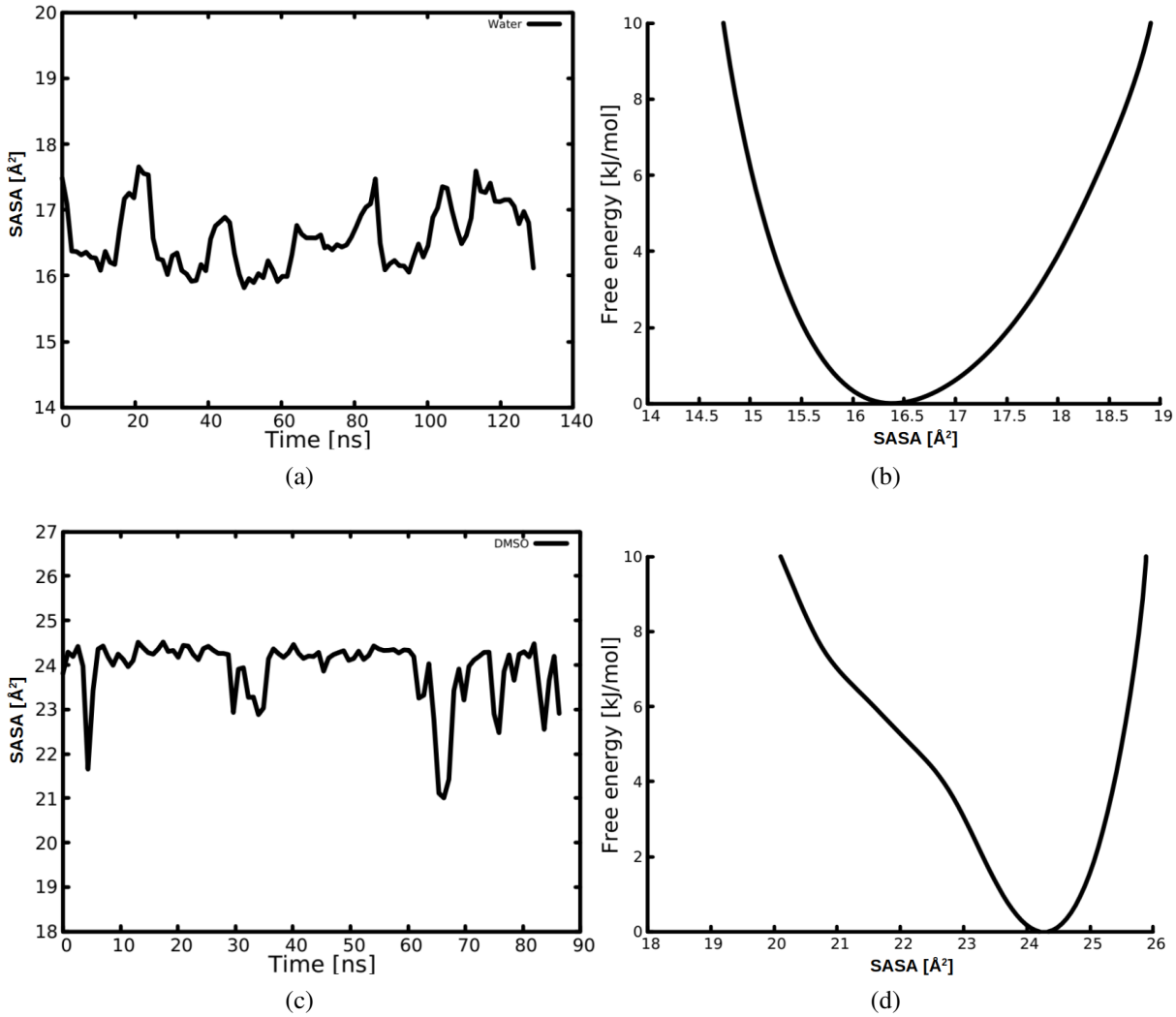


Figure 26: (a) and (c) are time dependent SASA values for chitosan nanoparticle in water and DMSO solvents, respectively. (b) and (d) free energy as the function of SASA in water and DMSO, respectively.

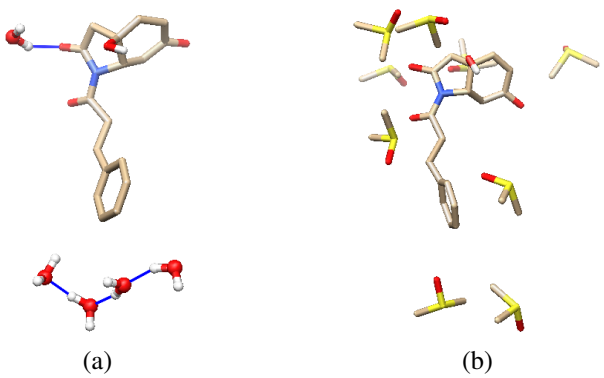


Figure 27: Representative snapshots showing the number of solvents in the first hydration shell for TouA bound to chitosan in (a) DMSO and in (b) water, chitosan is not shown for clarity.

4.3.2 Unbaised 1D free energy surfaces

The central idea of this subsection of the thesis is to understand the underlying effects of two solvents namely DMSO and water on the unbinding residence time and kinetics of TouA from chitosan nanoparticle. Free energies calculated from classical MD simulation and metadynamics are used to provide insight on the unbinding rates, furthermore, thermodynamic integration provides the related binding affinity. Before embarking to describe the residence time and kinetics on TouA unbinding, the thesis start by discussing the free energy surfaces obtained from both unbiased and biased simulations. Although classical MD simulation has limitations related to sampling problem, it provides the underlying picture of how the system behaves. In order to capture the unbinding behaviour in the two solvents, first the time dependent for TouA-chitosan minimum distance (d_{min}) over the simulation time of 140 ns in Fig. 28 are presented. The time dependent provides some information on the unbinding process, one can see that in DMSO a distance at 1.1-1.2 nm is mostly populated than in water where the distance at 0.35 nm most populates. In order to quantify the effect of solvents on the unbinding process the free energy (F) was calculated using the relation $F = -k_{\mathrm{B}}T\ln(P)$, where k_{B} is the Boltzmann constant, T is temperature and P is the probability. The free energy profiles presented in Fig. 28 suggest a clear solvents effect on TouA-chitosan interaction. The free energies in DMSO and water show a clear distinct free energy minima. The free energy for the unbinding state in DMSO indicates the presence of two minima at 0.35 and 1.19 nm corresponding to bound and unbound state, respectively. A more deeper minimum is observed for the unbound state as compared to the bound state which suggests that the complex in DMSO solvent exists more in an unbound state. However, in water a different behavior is observed, a deep free energy minimum was observed at 0.35 nm corresponding to the bound state. The free energy for unbound state was observed at a distance of 0.6 nm. The observed energy barrier for bound and unbound state in water is 2-fold higher than that in DMSO, this further explains the reason why there is large free energy in the unbound state in DMSO when compared to water. The differences in free energy in the two solvents is further explained by weak intermolecular interaction in DMSO than in water. The unbiased MD simulation provides an important feature on the system behaviour and suggests that the interaction in DMSO is poor which mostly favours the formation of a larger product as evidenced by the small energy barrier which disfavour the reaction going back to the reactants. Unlikely for water, the formation of the product is less favoured due to large energy barrier. In order to provide more information, the next subsection presents and discusses enhanced sampling results obtained using WT-MetaD which provide more insights on the unbinding residence time and kinetics of TouA in chitosan nanoparticle.

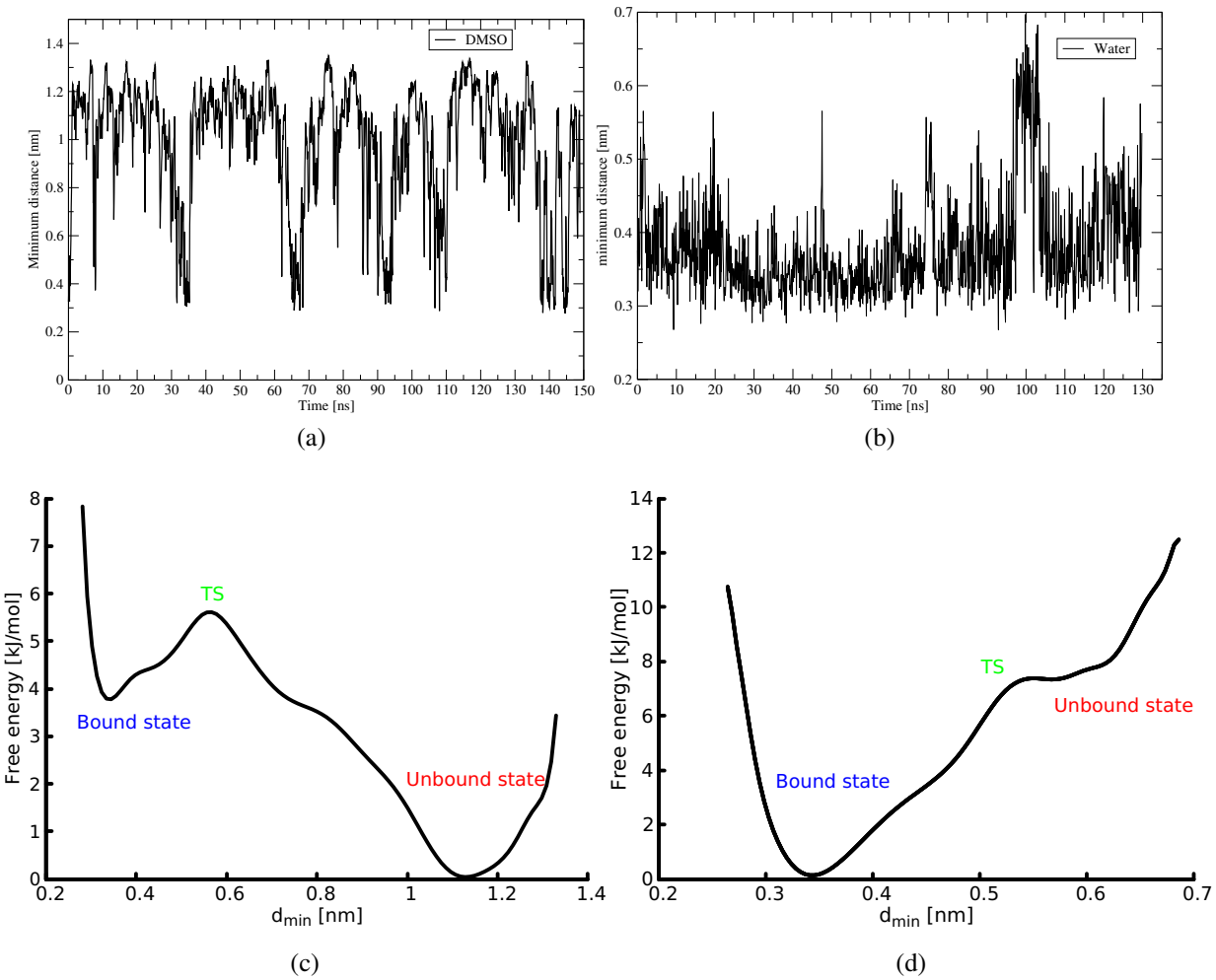


Figure 28: a-b, Time dependent for minimum distance between chitosan and TouA obtained from unbiased MD simulation during the 140 ns. Free energy profiles as the function of minimum distance between chitosan and TouA obtained from unbiased MD simulation, (c) in DMSO and (d) in water. TS is the transition state or barrier

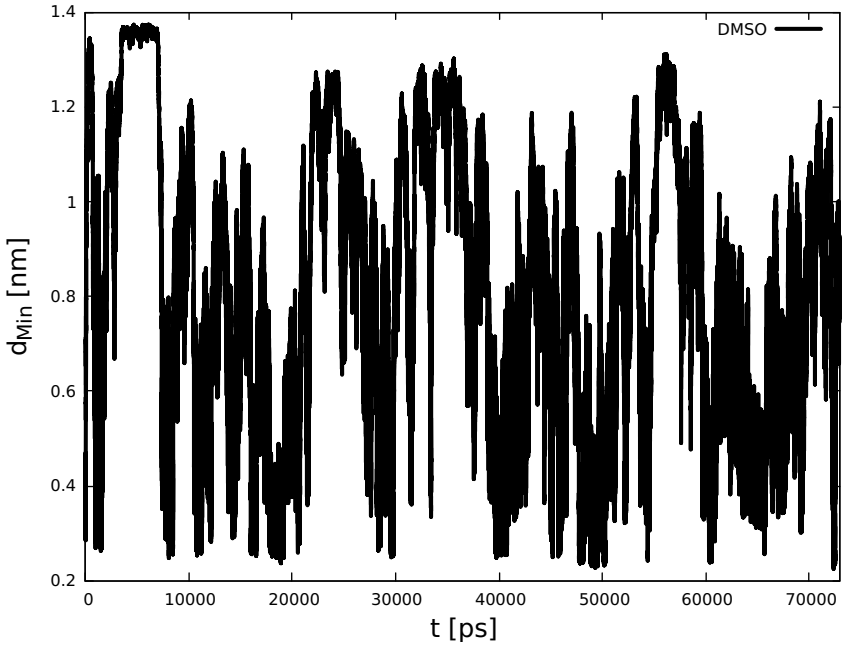
4.3.3 Baised 1D free energy surfaces

Because of the limitations of classical MD simulation in sampling rarely events and time constraints, solvent effects on drug unbinding kinetics and residence time are further investigated by WT-MetaD in water and DMSO. To describe the effects of solvents on the unbinding process, the following collective variables (CVs) are used: minimum distance (d_{min}) between specific groups of TouA defined in Fig. 25(a) and chitosan as CV5, coordination number between chitosan and TouA as CV6 and torsional angles as CV7. In order to describe the effects of solvents on the conformational changes of TouA bound to chitosan nanoparticle, the following CVs are used: keto-enol distance, end to end distance and torsional angles. The time-dependent of CV5 in the two solvents is presented in Fig. 29 which suggests system diffusivity with time implying convergence of the system. Then, the free energy surfaces to describe the unbinding process are calculated and discussed in the subsequent sections, first, the discussion begins by describing the 1D-free energy of the unbinding process.

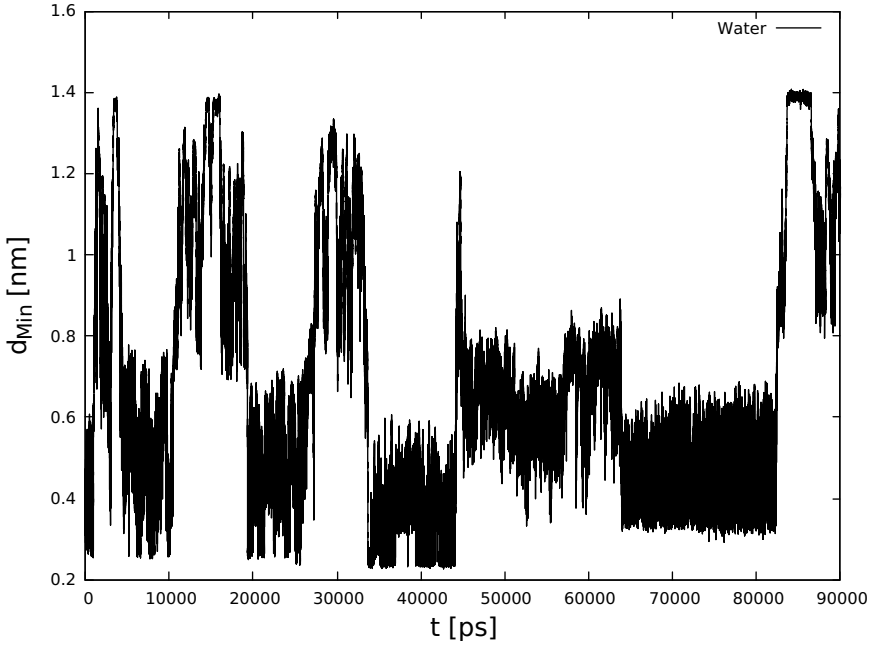
In Figure 30 the 1D free energy surface along the minimum distance obtained for water and DMSO systems are presented. The presented comparison helps to build the underlying intuition on the effects of the solvent on drug unbinding kinetics and residence time. Visual inspection on the 1D free energy indicates the two systems (water and DMSO) are different, suggesting that solvents affect the kinetics and residence time during the unbinding process in a different way. First, all systems show a similar binding behaviour characterized by a minimum distance at 0.3 nm which corresponds to electrostatic interaction between the chitosan and TouA groups. The effect of solvents is manifested by different free energies for the unbinding process as presented in Table 5 and Fig. 30, the free energy difference between the bound state and transition state (TS) in water is 6-folds higher to that in DMSO system. The smaller barrier of 5 kJ/mol in DMSO clearly shows that TouA would quickly be released, while in water it requires an amount of 34 kJ/mol to cross the barrier and hence slowly released. This is mostly advantageous for controlled drug delivery as it will take long time for the drug to be completely released. Based on the energy barrier one can see that there is a fast rate for the unbinding of TouA in DMSO than in water. Such differences could be brought by the difference in the electronegativity, solvent dielectric constant and polarity of the solvents. Water being a more polar solvent with higher dielectric constant results to formation of aggregated chitosan-TouA complex structure with strong intermolecular interaction while DMSO a polar aprotic solvent with lower dielectric constant tends to weaken the intermolecular hydrogen bonds formed between chitosan and TouA and hence resulting to a fast unbinding rate. The presented thermodynamic data for both water and DMSO systems integrate out the coordination number as the coordinate that would characterize the binding and unbinding kinetics process. Thus, the next subsection presents and discusses the 2D free energy surface obtained by biasing coordination number and minimum distance (d_{min}).

Table 5: Free energy differences (ΔF) between bound and TS (kJ/mol) as a function of minimum distance (d_{min}) for TouA binding in different solvents. ε is solvent dielectric constant and μ dipole moment (D). The values for dipole moment and dielectric (ε) constant are taken from Reichardt C., (2003).

Solvent	μ(D)	ε	ΔF
Water	1.86	80.1	34
DMSO	3.96	46.7	05



(a)



(b)

Figure 29: Time dependent of TouA unbinding process from chitosan nanoparticle in (a) DMSO and (b) Water.

4.3.4 2D free energy

The discussion begins by illustrating the 2D free energy surface obtained by biasing the minimum distance and coordination number as collective variables for both systems as shown in Appendix 1. The free energy surface for DMSO system reveals more information on the binding and unbinding kinetics characterized by different structural feature with four minima along

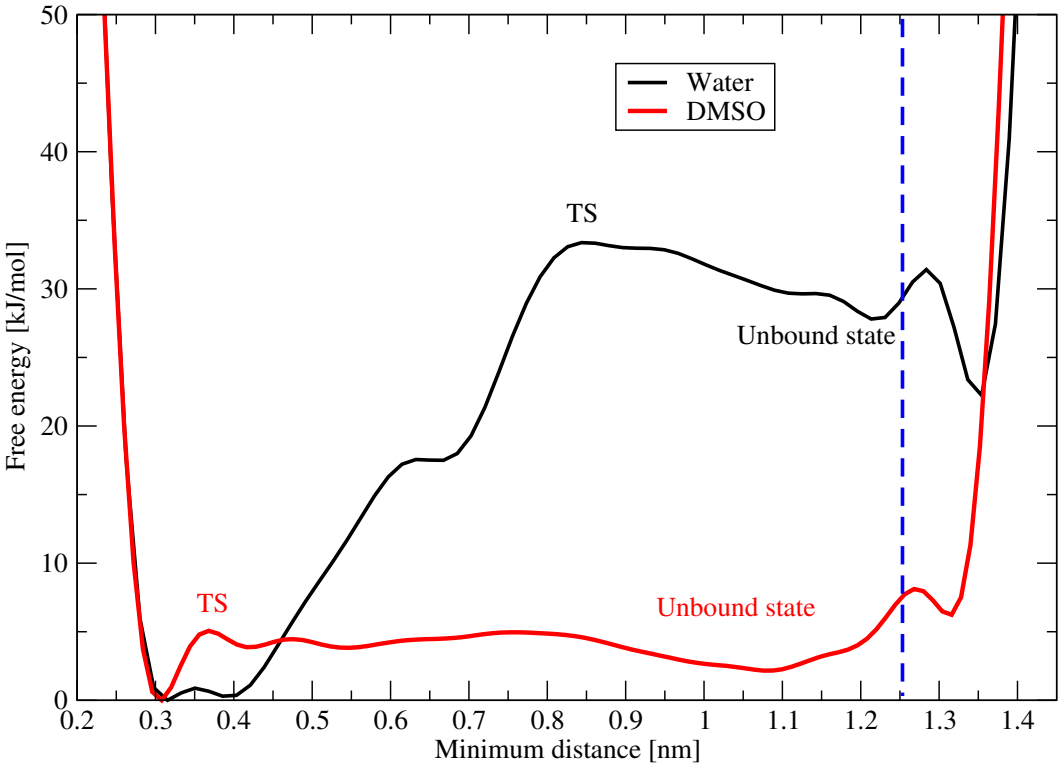


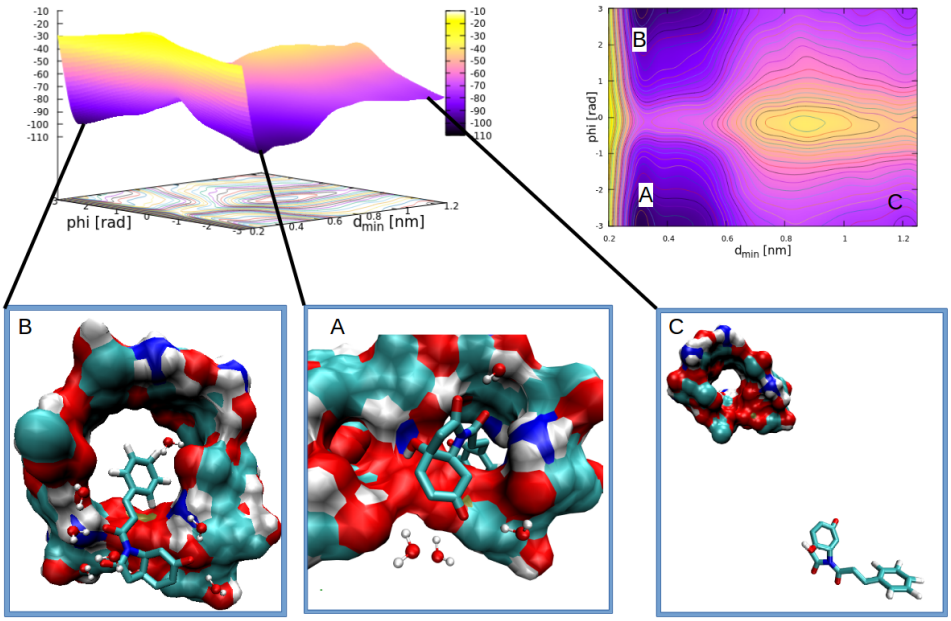
Figure 30: 1D free energy for the unbinding of TouA in water (black) and DMSO (red) as a function of CV minimum distance (d_{min}). The vertical dash line represents an artifact caused by the wall imposed during the simulation. TS is the transition state or barrier

the minimum distance and coordination number in different regions: As it has been observed before, the bound structure corresponds to a minimum distance of 0.3 nm with a coordination number of ≈ 3. The unbound state corresponds to a lower coordination number of less than 1. One can see that the bound and unbound minima are separated by other minima located at $d_{min} \approx 0.4$ and 0.55 nm with a coordination number of 0.5. Due to the small barrier as observed in Fig. 30, TouA in DMSO could escape from the bound state to unbound state more easily and hence released quickly from its carrier. This is also evidenced by the free energy for which has a different of 1 kJ/mol. However, for the system containing water solvent, a different behaviour is observed. One can see that the free energy for the bound state at $d_{min} \approx 0.3$ and 0.4 nm is large than the unbound state at $d_{min} \approx 1.2$ nm, suggesting that, water influences the intermolecular interaction of the two molecules by contributing the electrostatic interaction. Furthermore, the bound state and unbound state are separated by the energy barrier of ≈ 31 kJ/mol, causing TouA not to escape easily from its carrier.

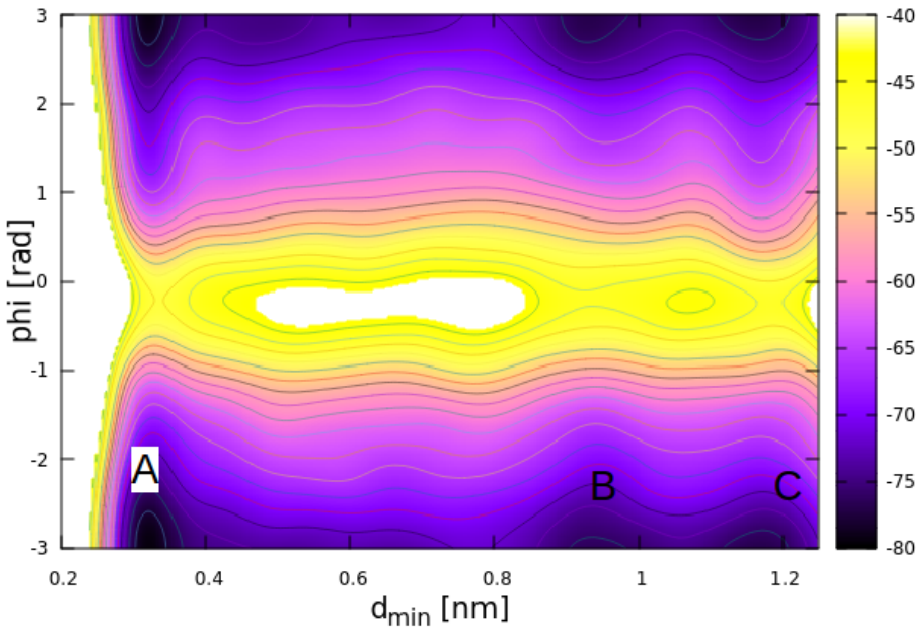
The change of chemical environment is well appreciated to have a greater effect on the dynamics, thermodynamics and functions of biologically important molecules (Levy *et al.*, 2001; Papadakis & Deligkiozi, 2019). The function and thermodynamics of binding between host and guest molecules can change when exposed to different solvents (Papadakis & Deligkiozi, 2019). This is evidenced by the free energy for the interaction of chitosan and TouA in DMSO

and water solvents. In the former solvent, the interaction tends to be weak when compared to the later solvent. This is attributed to the difference in polarity and electronegativity, as the DMSO is less electronegative and polar aprotic. On the other hand, the strong binding state for water is expected due to strong intermolecular interaction between water and chitosan/TouA. This is because water molecule interacts with the polar group of chitosan/TouA to form hydrogen bonding. Likely, the destabilization of hydrogen bonding in DMSO system results in weakening the interaction and hence TouA is released easier.

The choice of CVs is not a trivial process in metadynamics simulation, torsional angle (CV7) and minimum distance (CV5) were further used to describe the kinetic and unbinding process of TouA. Such CVs have been also used in protein-ligand unbinding kinetics experiment and provided useful insights into ligand unbinding (Limongelli *et al.*, 2013). In addition, the CVs are used to show the conformation change and motion of ligand as it unbinds. Using these CVs one can see that the system in water visits the bound state several times with a $d_{min} < 0.6$ nm and torsional angle $\phi = \pm 3$ rad. The corresponding free energy and associated structures to aid the discussion are presented in Fig. 31. Two energy basins A and B are observed for bound state, however, basin B is the preparation state for TouA to unbind from chitosan. At this basin TouA orients in a such way it starts to be expelled out from chitosan. Another important observation from basin B is that, at this state chitosan also undergoes conformation adaptability which favour TouA to be released.It is worth to note that the conformational adaptability of chitosan at basin B is different to that in A. In the unbound state basin C, TouA has no contacts with its carrier and has several orientations and chitosan retain its conformation similar to that in basin B. In DMSO, a completely reverse process where the unbound basin C is observed with large free energy than the bound basin A. An intermediate state for unbinding is observed at basin B before TouA is completely unbound and full solvated. In a closer observation, solvent effects on the orientation of TouA is well manifested in the bound and unbound basins in the two solvents. In water the orientation with large free energy at $\phi = \pi/2$ and π rad, while in DMSO both bound and unbound basins shows the rotation at $\phi = \pi$ rad was observed. Visual inspection to the snapshots show that the functionalized TPP groups play an important role in the chitosan-TouA dissociation mechanism as the release pathway of TouA proceeds near the TPP groups.



(a)



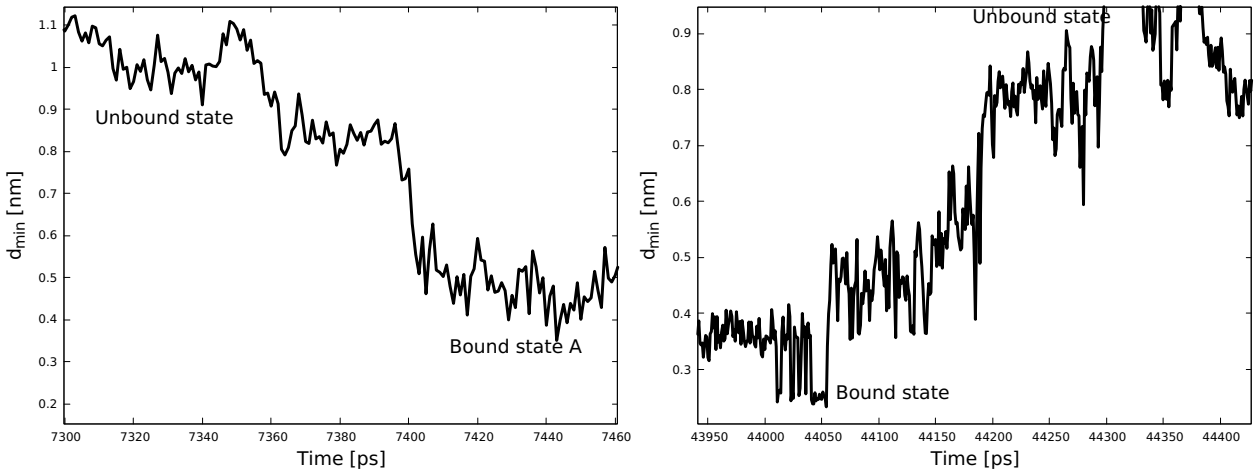
(b)

Figure 31: (a) Binding and unbinding processes of TouA in water. Atoms are shown by their colour: Hydrogen is white, Oxygen is red, Nitrogen is blue and Carbon is light blue. (b) the binding conformation corresponds to A while unbinding conformation are labled B and C in DMSO.

4.3.5 Drug kinetics and residence time (τ_{AB})

As described earlier, drug kinetics and residence time play an important role in drug design. The residence time defined as the reciprocal of the dissociation rate constant ($\tau = 1/k_{\text{off}}$) is known to depend on conformations stabilization. In this work the time taken by TouA from the bound to unbound state is quantified,

As presented in Fig. 37, the free energy for the bound and unbound states is related to the activation barrier for the host-guest dissociation, which describes the rate of the reaction. Higher activation barrier results to slow reaction rates. The time exchange from a bound to unbound state (dissociation) is an important parameter to describe the rate of the reaction. In this work, τ_{ex} is the time exchange for TouA from state A (bound) to B (unbound) and has an average value of $\approx$ 100 ps as shown in Fig. 32.



(a) (b)

Figure 32: Time exchange τ_{ex} from bound to unbound state in (a) DMSO and (b) water.

Table 6 shows the results for τ_{AB} in different solvents. One can see that there is solvent effect on the kinetics and residence time of TouA to unbound from chitosan. In water, TouA shows a slow rate of $k_{\text{off}} = 0.045/\mu$s compared to DMSO where the k_{off} is fast 1000 times. The k_{off} values suggest that TouA in presence of DMSO is expelled out from chitosan faster and has weak interaction than in water. As previously reported (Levy *et al.*, 2001; Papadakis & Deligkiozi, 2019) that solvent affects the kinetics and supramolecular host-guest interaction, the current results suggest that the interaction of chitosan and TouA is highly solvent dependent. To attain an optimal release of TouA from chitosan, this study suggests that the formulation should be done in water and not DMSO. Until the time of writing this thesis, there isn't any experimental work which reports the experimental formulation of chitosan-TouA in DMSO. However, the formulation of chitosan-TouA in water has been previously reported (Rwegasila *et al.*, 2018). The computational results on drug kinetics and residence time in water reported in this thesis can be compared to the experimental work, where the release kinetics of TouA was sustained for 72 hrs similar to the obtained slow rate, suggesting a slow release of TouA in water. Similarly, the experimental encapsulation efficiency and loading capacity are also compared to free energy landscape where in water a large free energy in a bound state was observed which supports strong supramolecular interaction. The reported findings are useful and provide a piece of baseline information on how one can decide to choose a solvent for supramolecular interaction.

Table 6: Residence time (τ_{AB}) and kinetics rate (k_{off}) for unbinding process of TouA in two solvents.

System	$\tau_{AB}(\mu s)$	k_{off} /μs
water	22	0.045
DMSO	0.001	1000

4.3.6 Conformation of TouA bound to chitosan in solvents

Furthermore, the kinetic and residence time of TouA to its stability in the two solvents are linked. In order investigate the stability the following CVs were biased: keto-enol distance formed by atoms 11 and enolic group 12, end to end distance formed by atoms 7 and 21 and spacer dihedral angles formed by atoms 8, 14, 16, 17 to determine internal changes in TouA molecule. Figure 33 shows the 1D free energy as the function of keto-enol distance. One can see that, the keto-enol distance shows a minimum at $d_{min} \approx 0.35$ and 0.55 nm suggesting that TouA exists with a conformation of enolic hydrogen pointing towards the keto oxygen and away from the keto oxygen as shown in Fig. 34c. The distance at 0.35 nm suggests an intramolecular interaction of TouA molecule while the distance at 0.55 nm suggests an intermolecular interaction with chitosan or solvent as shown in an insert in Fig. 33. Table 7 further shows the free energy difference (ΔF) for keto-enol distances in the two solvents. In both solvents one can observe a conformation at a distance of 0.55 nm to be dominated, suggesting that, although TouA may form intramolecular hydrogen bonding, it exists in a stable conformation with the enolic hydrogen pointing away from the keto oxygen. However, comparing the two free energies in water and DMSO one can see large free energy in water. This behaviour is expected as a result of higher electronegativity tendency of water molecule which tends to form intermolecular interaction with the enolic hydrogen of TouA causing a C-O bond rotation (see an insert in Fig. 33). To capture more information on the TouA kinetics, Fig. 34 shows the 2D free energy as a function of CV3 and CV4. Different behaviours of TouA in the two solvents is portrayed. In DMSO TouA shows two minima for the keto-enol distance while the end to end distance shows a stable minimum at ≈ 1.2 nm suggesting that TouA in DMSO exists more in stretched conformation. However, this is not the case in water where two distinct minima corresponding for bent and stretched conformation are observed. For further discuss, as an example, the associated structures at each state in Fig. 34c are presented. Both 1D and 2D free energies in two solvents portray a different behaviour and stability, where in water TouA is 2-folds thermodynamic stable than in DMSO. The measured dihedral angles and end to end distances are presented in Fig. 35, the dihedral angles ϕ further show a rotational at $\phi = \pm 3$ rad with a more stretched conformation in DMSO and both bent and stretched conformation in water. The rotation at $\phi = \pm 3$ suggests the *trans*-conformation to be more thermodynamic stable than *cis*-conformation.

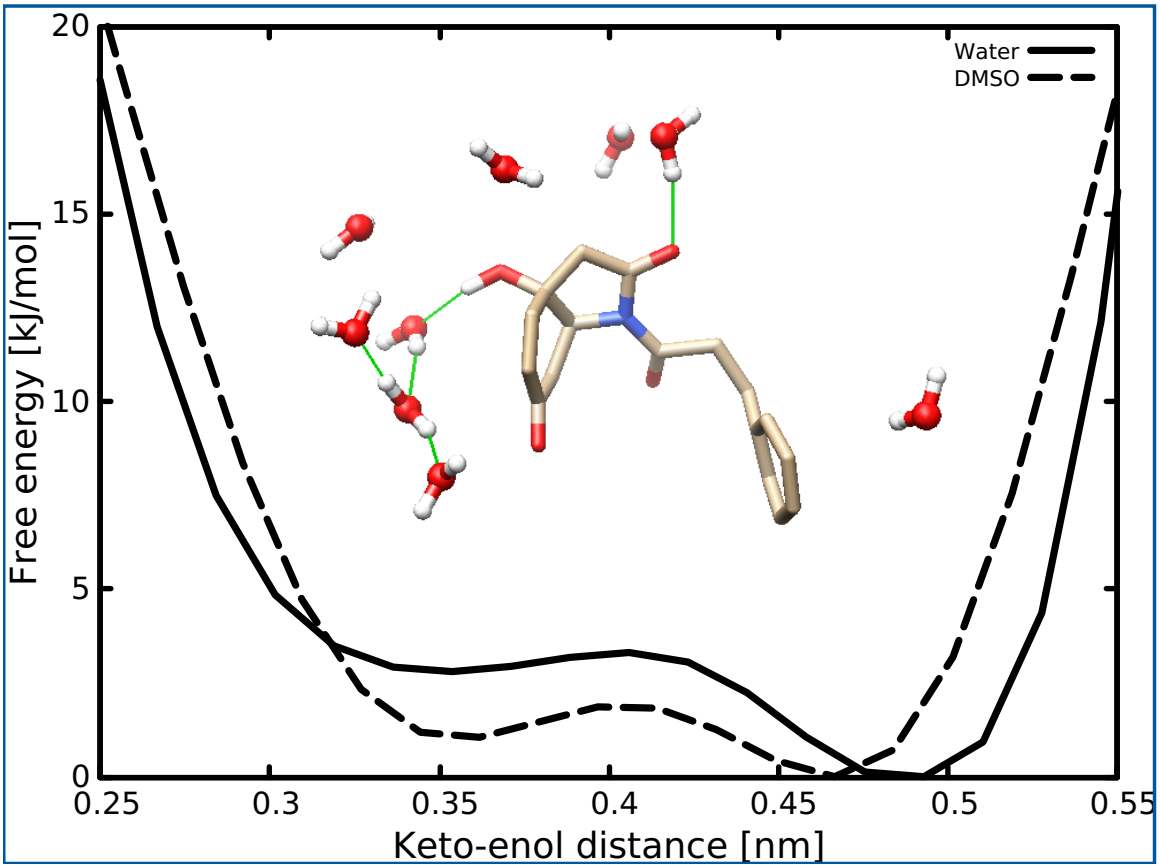
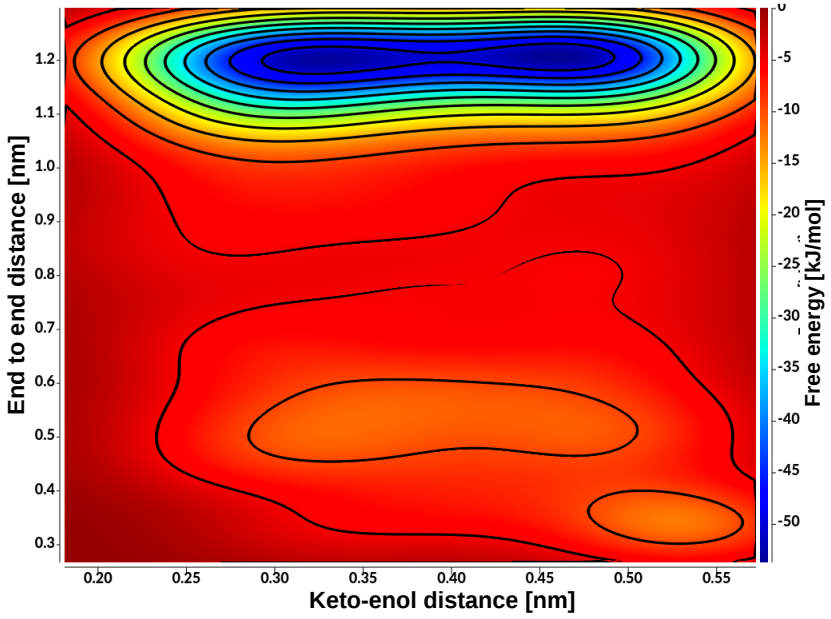


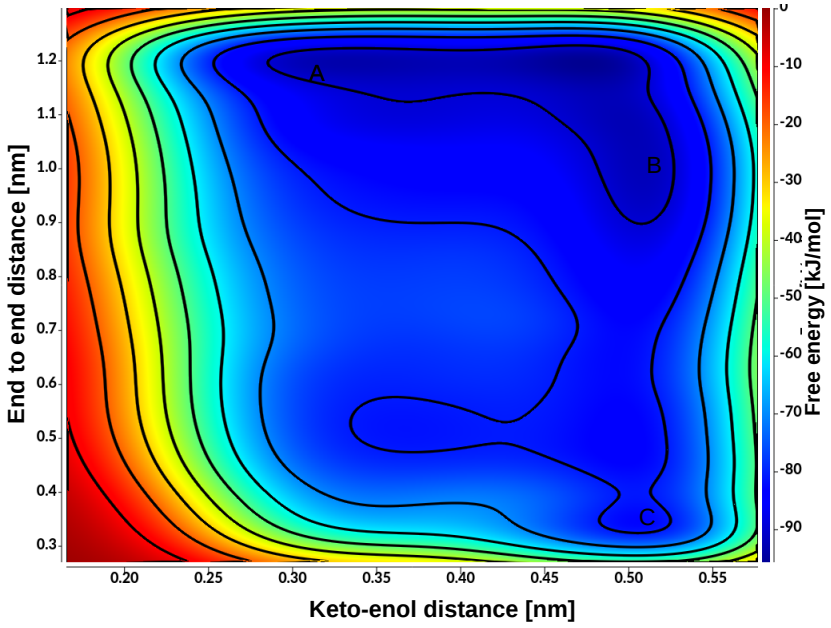
Figure 33: 1D free energy on solvent effects on TouA keto-enol distance when bound to chitosan. An insert shows the intermolecular interaction between TouA and water.

Table 7: Free energy difference (ΔF) (kJ/mol) for keto-enol distances of TouA bound to chitosan in two solvent.

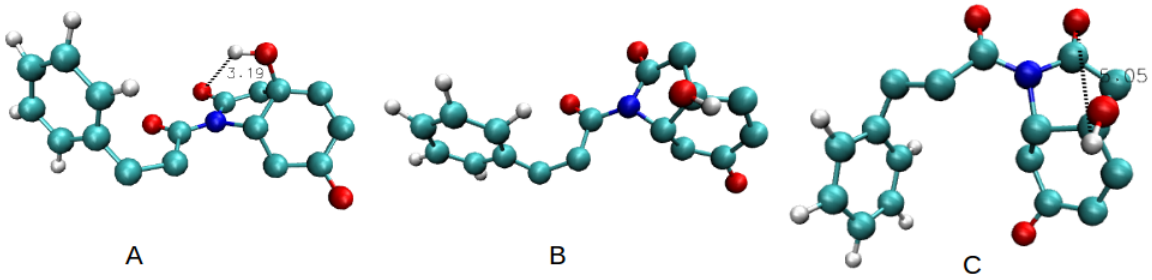
System	ΔF
water	2.83
DMSO	1.05



(a)

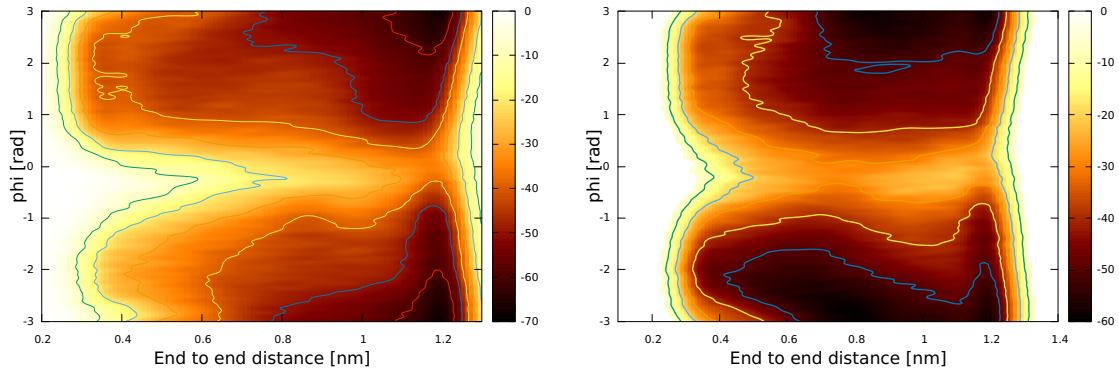


(b)



(c)

Figure 34: a-b, 2D free energy of TouA conformation when bound to chitosan (a) in DMSO and (b) in water. (c) is the conformation of TouA bound to chitosan in water taken at the following coordinates A (0.319,1.19), B (0.51,0.95) and C (0.509,0.386).



(a) (b)

Figure 35: Orientations of TouA bound to chitosan in two solvents. (a) in DMSO and (b) in water.

4.3.7 Thermodynamic stability of chitosan-TouA complex

Previous subsections have explored the solvent related effects on kinetics and residence time of chitosan-TouA complex in DMSO and water. This subsection describes the solvent effects on the thermodynamic stability of the complex in solvents. Previously, it has been shown that, changes in solvent polarity has a drastic effects on the thermodynamics stability of the complex (Papadakis & Deligkiozi, 2019). The thermodynamic stability of chitosan-TouA complex in two solvents is further investigated by means of TI free energy. Although it is computationally expensive, TI is an effective method for calculating the binding free energy of host-guest molecules in solvents. The calculated TI binding free energy presented in Table 8 further shows the related solvent effects on thermodynamics stability of the complex. The thermodynamics stability for chitosan-TouA are different in these two solvents where in water more stable complex was observed than in DMSO. This is because highly polar solvents such as water have major effects on supramolecular host-guest complexes (Papadakis & Deligkiozi, 2019). These result suggests that, although water results to thermodynamics stable complex, DMSO has a reverse effect. Unfortunately, the favourable interaction in DMSO was not obtained as the free energy was $>$ 30 kcal/mol. Reverse effect of solvents on supramolecular complexes has also been previously observed for different solvents (Kang & Rebek Jr, 1996). In this thesis, both metadynamics and TI deciphered information on the interaction of chitosan-TouA in DMSO to be more weaker and unfavourable when compared to water where there is favourable and strong interaction.

Generally, the preceding discussion has demonstrated the effects of solvents on biomolecular host-guest kinetics and residence time. Two solvents DMSO and water were used to study the influence of solvents. It was observed that more polar protic solvent results to more stabilization of host-guest interactions and polar aprotic solvents tend to have reverse effects on the binding process favouring high unbinding rates. This influence can also be attributed to the nature of

the encapsulated drug molecules. A similar approach can be extended to investigate host-guest interaction for different nanoparticle and drugs.

Table 8: Thermodynamic integration (TI) binding free energy for chitosan-TouA complex in DMSO and water.

System	$\Delta G_{binding}$ kcal/mol
DMSO	poor bound (>30)
Water	-10.83 ± 0.33

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

This thesis has investigated the role of solvents in conformational fluctuations and drug interaction with an implication for drug design and discovery. The thesis has explored various docking protocols used in drug design and discovery, the following conclusions are drawn from the study:

(i) There is a sensitivity of results to different docking protocols. The use of ensemble structures with and without water reduced the binding energies compared to crystal structures.

(ii) Holo ensembles with ligands that bound strongly increase the binding energy close to crystal structures. The use of ensemble structures holo or apo can improve the binding energy, but this is not always the case. The docking protocols reported here pave the way on drug design and suggest an integration of such protocols for various protein-ligand complex.

(iii) Using metadynamics, an enhanced sampling approach, the thesis also investigated the effects of different solvents on the stability, conformation and orientation of natural products using curcumin as a model molecule in polar protic, polar aprotic and non-polar solvents. The conformation and stability of curcumin is solvent dependent. It is further observed that stability and conformation of curcumin depend on solvent polarity.

(iv) Solvents play a major role in determining molecules conformation, stability and biological functions. To further ascertain this thesis investigated solvent effects on host-guest kinetics and residence time using chitosan-TouA in DMSO and water as model systems. High unbinding kinetics rate (k_{off}) in DMSO and slow unbinding kinetics (k_{off}) in water were observed. TouA was observed to be released more quickly in DMSO than in water.

(v) The approaches used in this study *viz:* different docking protocols, role of solvents on drug interactions with protein and nanoparticles (macromolecules) are recommended for applications in various drug design targeting several indications.

5.2 Recommendations

The following recommendations are drawn from the study:

(i) The small molecules reported in this thesis as potential inhibitors for Hsp90 posses known pharmacological profiles and are used for treatment of other indications. These molecules are highly recommended for further pre-clinical and clinical studies as cancer chemotherapeutic agents targeting Hsp90.

(ii) The exploration of natural product crystals heterogeneity interaction with solvents such as water is highly recommended for a deeper understanding of the behaviour of water near the natural products/drug crystal surface. This will further provide more understanding of the issues of solubility and stability of molecules in different solvents. Studies leading to the investigation of heterogeneity at curcumin crystal water interface is highly recommended for better understanding of the origin of poorly water solubility of curcumin.

(iii) Investigation of the crystal growth and polymorphism changes of the investigated small molecules is highly recommended to establish their related stability.

REFERENCES

Abbasi, M., Sadeghi-Aliabadi, H., & Amanlou, M. (2018). 3d-qsar, molecular docking, and molecular dynamic simulations for prediction of new hsp90 inhibitors based on isoxazole scaffold. *Journal of Biomolecular Structure and Dynamics*, *36*(6), 1463–1478.

Abraham, M., van der Spoel, D., Lindahl, E., & Hess, B. (2015). The gromacs development team gromacs user manual version 5.1. 2; 2016. *SoftwareX*, *1*, 19.

Aldeghi, M., Bluck, J. P., & Biggin, P. C. (2018). Absolute alchemical free energy calculations for ligand binding: A beginners guide. In *Computational Drug Discovery and Design*. Springer.

Aldeghi, M., Heifetz, A., Bodkin, M. J., Knapp, S., & Biggin, P. C. (2016). Accurate calculation of the absolute free energy of binding for drug molecules. *Chemical Science*, *7*(1), 207–218.

Amaro, R. E., Baron, R., & McCammon, J. A. (2008a). An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *Journal of Computer-Aided Molecular Design*, *22*(9), 693–705.

Amaro, R. E., Schnaufer, A., Interthal, H., Hol, W., Stuart, K. D., & McCammon, J. A. (2008b). Discovery of drug-like inhibitors of an essential rna-editing ligase in trypanosoma brucei. *Proceedings of the National Academy of Sciences*, *105*(45), 17278–17283.

Austin, C., Pettit, S. N., Magnolo, S. K., Sanvoisin, J., Chen, W., Wood, S. P., Freeman, L. D., Pengelly, R. J., & Hughes, D. E. (2012). Fragment screening using capillary electrophoresis (cefrag) for hit identification of heat shock protein 90 atpase inhibitors. *Journal of Biomolecular Screening*, *17*(7), 868–876.

Barakat, K., Mane, J., Friesen, D., & Tuszynski, J. (2010). Ensemble-based virtual screening reveals dual-inhibitors for the p53–mdm2/mdmx interactions. *Journal of Molecular Graphics and Modelling*, *28*(6), 555–568.

Barakat, K., & Tuszynski, J. (2011). Relaxed complex scheme suggests novel inhibitors for the lyase activity of dna polymerase beta. *Journal of Molecular Graphics and Modelling*, *29*(5), 702–716.

Barducci, A., Bussi, G., & Parrinello, M. (2008). Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Physical Review Letters*, *100*(2), 020603.

Barker, J. J., Barker, O., Courtney, S. M., Gardiner, M., Hesterkamp, T., Ichihara, O., Mather, O., Montalbetti, C. A., Müller, A., & Varasi, M. (2010). Discovery of a novel hsp90 inhibitor by fragment linking. *Chemmedchem*, *5*(10), 1697–1700.

Baum, B., Mohamed, M., Zayed, M., Gerlach, C., Heine, A., Hangauer, D., & Klebe, G. (2009). More than a simple lipophilic contact: a detailed thermodynamic analysis of nonbasic residues in the s1 pocket of thrombin. *Journal of Molecular Biology*, *390*(1), 56–69.

Baum, B., Muley, L., Smolinski, M., Heine, A., Hangauer, D., & Klebe, G. (2010). Non-additivity of functional group contributions in protein–ligand binding: a comprehensive study by crystallography and isothermal titration calorimetry. *Journal of Molecular Biology*, *397*(4), 1042–1054.

Berendsen, H. J., Postma, J. V., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, *81*(8), 3684–3690.

Bhat, R., Tummalapalli, S. R., & Rotella, D. P. (2014). Progress in the discovery and development of heat shock protein 90 (hsp90) inhibitors: Miniperspective. *Journal of Medicinal Chemistry*, *57*(21), 8718–8728.

Bissantz, C., Kuhn, B., & Stahl, M. (2010). A medicinal chemists guide to molecular interactions. *Journal of Medicinal Chemistry*, *53*(14), 5061–5084.

Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., & Broglia, R. A. (2009). Plumed: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications*, *180*(10), 1961–1972.

Brice, A. R., & Dominy, B. N. (2011). Analyzing the robustness of the mm/pbsa free energy calculation method: application to dna conformational transitions. *Journal of Computational Chemistry*, *32*(7), 1431–1440.

Brooijmans, N. (2009). Docking methods, ligand design and validating data sets in the structural genomics era. *John Wiley & Sons Inc.: New York, NY, USA*.

Bussi, G., Donadio, D., & Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, *126*(1), 014101.

Carr, A. C., Piunova, V. A., Maarof, H., Rice, J. E., & Swope, W. C. (2018). Influence of solvent on the drug-loading process of amphiphilic nanogel star polymers. *The Journal of Physical Chemistry B*, *122*(21), 5356–5367.

Cavasotto, C. N., Orry, W., & Andrew, J. (2007). Ligand docking and structure-based virtual screening in drug discovery. *Current Topics in Medicinal Chemistry*, *7*(10), 1006–1014.

Cavenagh, J., Oakervee, H., Baetiong-Caguioa, P., Davies, F., Gharibo, M., Rabin, N., Kurman, M., Novak, B., Shiraishi, N., & Nakashima, D. (2017). A phase i/ii study of kw-2478, an hsp90 inhibitor, in combination with bortezomib in patients with relapsed/refractory multiple myeloma. *British Journal of Cancer*, *117*(9), 1295.

Chan, A. H., Wereszczynski, J., Amer, B. R., Yi, S. W., Jung, M. E., McCammon, J. A., & Clubb, R. T. (2013). Discovery of s taphylococcus aureus sortase a inhibitors using virtual screening and the relaxed complex scheme. *Chemical Biology & Drug Design*, *82*(4), 418–428.

Chatterjee, S., & Burns, T. (2017). Targeting heat shock proteins in cancer: a promising therapeutic approach. *International Journal of Molecular Sciences*, *18*(9), 1978.

Cheatham, T. I., Miller, J., Fox, T., Darden, T., & Kollman, P. (1995). Molecular dynamics simulations on solvated biomolecular systems: the particle mesh ewald method leads to stable trajectories of dna, rna, and proteins. *Journal of the American Chemical Society*, *117*(14), 4193–4194.

Chen, F., Liu, H., Sun, H., Pan, P., Li, Y., Li, D., & Hou, T. (2016). Assessing the performance of the mm/pbsa and mm/gbsa methods. 6. capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking. *Physical Chemistry Chemical Physics*, *18*(32), 22129–22139.

Cheng, L. S., Amaro, R. E., Xu, D., Li, W. W., Arzberger, P. W., & McCammon, J. A. (2008). Ensemble-based virtual screening reveals potential novel antiviral compounds for avian influenza neuraminidase. *Journal of Medicinal Chemistry*, *51*(13), 3878–3894.

Chiosis, G. (2006). Discovery and development of purine-scaffold hsp90 inhibitors. *Current Topics in Medicinal Chemistry*, *6*(11), 1183–1191.

Comeau, S. R., Gatchell, D. W., Vajda, S., & Camacho, C. J. (2004). Cluspro: a fully automated algorithm for protein–protein docking. *Nucleic Acids Research*, *32*(2), 96–99.

Cui, D., Zhang, B. W., Matubayasi, N., & Levy, R. M. (2018). The role of interfacial water in protein–ligand binding: Insights from the indirect solvent mediated potential of mean force. *Journal of Chemical Theory and Computation*, *14*(2), 512–526.

da Silva, G. C., Cardozo, T. M., Amarante, G. W., Abreu, C. R., & Horta, B. A. (2018). Solvent effects on the decarboxylation of trichloroacetic acid: insights from ab initio molecular dynamics simulations. *Physical Chemistry Chemical Physics*, *20*(34), 21988–21998.

Darden, T., York, D., & Pedersen, L. (1993). Particle mesh ewald: An n log (n) method for ewald sums in large systems. *The Journal of Chemical Physics*, *98*(12), 10089–10092.

David, L., Nielsen, P. A., Hedstrom, M., & Norden, B. (2005). Scope and limitation of ligand docking: Methods, scoring functions and protein targets. *Current Computer-Aided Drug Design*, *1*(3), 275–306.

de Graaf, C., Oostenbrink, C., Keizers, P. H., van der Wijst, T., Jongejan, A., & Vermeulen, N. P. (2006). Catalytic site prediction and virtual screening of cytochrome p450 2d6 substrates by consideration of water and rescoring in automated docking. *Journal of Medicinal Chemistry*, *49*(8), 2417–2430.

De Vivo, M., Masetti, M., Bottegoni, G., & Cavalli, A. (2016). Role of molecular dynamics and related methods in drug discovery. *Journal of Medicinal Chemistry*, *59*(9), 4035–4061.

Dias, R., de Azevedo, J., & Walter, F. (2008). Molecular docking algorithms. *Current Drug Targets*, *9*(12), 1040–1047.

Dominguez, C., Boelens, R., & Bonvin, A. M. (2003). Haddock: a protein- protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*(7), 1731–1737.

Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*, *9*(1), 71.

Eachkoti, R., Reddy, M. R., Lieu, Y. K., Cosenza, S. C., & Reddy, E. P. (2014). Identification and characterisation of a novel heat shock protein 90 inhibitor ono4140. *European Journal of Cancer*, *50*(11), 1982–1992.

Ernst, J. T., Neubert, T., Liu, M., Sperry, S., Zuccola, H., Turnbull, A., Fleck, B., Kargo, W., Woody, L., & Chiang, P. (2014). Identification of novel hsp90α/β isoform selective inhibitors using structure-based drug design. demonstration of potential utility in treating cns disorders such as huntingtons disease. *Journal of Medicinal Chemistry*, *57*(8), 3382–3400.

Franca, E. F., Freitas, L. C., & Lins, R. D. (2011). Chitosan molecular structure as a function of n-acetylation. *Biopolymers*, *95*(7), 448–460.

Fu, I. W., Markegard, C. B., & Nguyen, H. D. (2014). Solvent effects on kinetic mechanisms of self-assembly by peptide amphiphiles via molecular dynamics simulations. *Langmuir*, *31*(1), 315–324.

Ge, J., Normant, E., Porter, J. R., Ali, J. A., Dembski, M. S., Gao, Y., Georges, A. T., Grenier, L., Pak, R. H., & Patterson, J. (2006). Design, synthesis, and biological evaluation of hydroquinone derivatives of 17-amino-17-demethoxygeldanamycin as potent, water-soluble inhibitors of hsp90. *Journal of Medicinal Chemistry*, *49*(15), 4606–4615.

Genheden, S., & Ryde, U. (2015). The mm/pbsa and mm/gbsa methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, *10*(5), 449–461.

Gereben, O., & Pusztai, L. (2011). On the accurate calculation of the dielectric constant from molecular dynamics simulations: The case of spc/e and swm4-dp water. *Chemical Physics Letters*, *507*(1-3), 80–83.

Gewirth, D. T. (2016). Paralog specific hsp90 inhibitors–a brief history and a bright future. *Current Topics in Medicinal Chemistry*, *16*(25), 2779.

Grubmüller, H. (1995). Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Physical Review E*, *52*(3), 2893.

Gupta, S. D., Bommaka, M. K., Mazaira, G. I., Galigniana, M. D., Subrahmanyam, C. V. S., Gowrishankar, N. L., & Raghavendra, N. M. (2015). Molecular docking study, synthesis and biological evaluation of mannich bases as hsp90 inhibitors. *International Journal of Biological Macromolecules*, *80*, 253–259.

Gupta, S. D., Snigdha, D., Mazaira, G. I., Galigniana, M. D., Subrahmanyam, C., Gowrishankar, N., & Raghavendra, N. (2014). Molecular docking study, synthesis and biological evaluation of schiff bases as hsp90 inhibitors. *Biomedicine & Pharmacotherapy*, *68*(3), 369–376.

Hani, U., & Shivakumar, H. (2014). Solubility enhancement and delivery systems of curcumin a herbal medicine: A review. *Current Drug Dselivery*, *11*(6), 792–804.

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, *4*(1), 17.

Hazra, M. K., Roy, S., & Bagchi, B. (2014). Hydrophobic hydration driven self-assembly of curcumin in water: Similarities to nucleation and growth under large metastability, and an analysis of water dynamics at heterogeneous surfaces. *The Journal of Chemical Physics*, *141*(18), 18C501.

Hernández-Rodríguez, M., Correa-Basurto, J., Gutiérrez, A., Vitorica, J., & Rosales-Hernández, M. C. (2016). Asp32 and asp228 determine the selective inhibition of bace1 as shown by docking and molecular dynamics simulations. *European Journal of Medicinal Chemistry*, *124*, 1142–1154.

Hertlein, E., Wagner, A. J., Jones, J., Lin, T. S., Maddocks, K. J., Towns, W. H., Goettl, V. M., Zhang, X., Jarjoura, D., & Raymond, C. A. (2010). 17-dmag targets the nuclear factor-κb family of proteins to induce apoptosis in chronic lymphocytic leukemia: clinical implications of hsp90 inhibition. *Blood*, *116*(1), 45–53.

Hess, B., Bekker, H., Berendsen, H. J., & Fraaije, J. G. (1997). Lincs: a linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, *18*(12), 1463–1472.

Hevener, K. E., Zhao, W., Ball, D. M., Babaoglu, K., Qi, J., White, S. W., & Lee, R. E. (2009). Validation of molecular docking programs for virtual screening against dihydropteroate synthase. *Journal of Chemical Information and Modeling*, *49*(2), 444–460.

Horn, H. W., Swope, W. C., Pitera, J. W., Madura, J. D., Dick, T. J., Hura, G. L., & Head-Gordon, T. (2004). Development of an improved four-site water model for biomolecular simulations: Tip4p-ew. *The Journal of Chemical Physics*, *120*(20), 9665–9678.

Hou, T., Wang, J., Li, Y., & Wang, W. (2010). Assessing the performance of the mm/pbsa and mm/gbsa methods. 1. the accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling*, *51*(1), 69–82.

Huang, W., Blinov, N., Wishart, D. S., & Kovalenko, A. (2015). Role of water in ligand binding to maltose-binding protein: insight from a new docking protocol based on the 3d-rism-kh molecular theory of solvation. *Journal of Chemical Information and Modeling*, *55*(2), 317–328.

Ignjatović, M. M., Caldararu, O., Dong, G., Munoz-Gutierrez, C., Adasme-Carreno, F., & Ryde, U. (2016). Binding-affinity predictions of hsp90 in the d3r grand challenge 2015 with docking, mm/gbsa, qm/mm, and free-energy simulations. *Journal of Computer-Aided Molecular Design*, *30*(9), 707–730.

Ilnytskyi, J., Patsahan, T., & Pizio, O. (2016). On the properties of the curcumin molecule in water. exploration of the opls-united atom model by molecular dynamics computer simulation. *Journal of Molecular Liquids*, *223*, 707–715.

Immormino, R. M., Kang, Y., Chiosis, G., & Gewirth, D. T. (2006). Structural and quantum chemical studies of 8-aryl-sulfanyl adenine class hsp90 inhibitors. *Journal of Medicinal Chemistry*, *49*(16), 4953–4960.

Izatt, R. M., Pawlak, K., Bradshaw, J. S., & Bruening, R. L. (1995). Thermodynamic and kinetic data for macrocycle interaction with cations, anions, and neutral molecules. *Chemical Reviews*, *95*(7), 2529–2586.

Jain, A. N. (2003). Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of Medicinal Chemistry*, *46*(4), 499–511.

Jarzynski, C. (1997). Nonequilibrium equality for free energy differences. *Physical Review Letters*, *78*(14), 2690.

Jiang, F., Wang, H. J., Jin, Y. H., Zhang, Q., Wang, Z. H., Jia, J. M., Liu, F., Wang, L., Bao, Q. C., & Li, D. D. (2016). Novel tetrahydropyrido [4, 3-d] pyrimidines as potent inhibitors of chaperone heat shock protein 90. *Journal of Medicinal Chemistry*, *59*(23), 10498–10519.

Jong, K., Grisanti, L., & Hassanali, A. (2017). Hydrogen bond networks and hydrophobic effects in the amyloid β30–35 chain in water: A molecular dynamics study. *Journal of Chemical Information and Modeling*, *57*(7), 1548–1562.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *79*(2), 926–935.

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, *118*(45), 11225–11236.

Jorgensen, W. L., & Tirado-Rives, J. (1988). The opls [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, *110*(6), 1657–1666.

Kang, J., & Rebek Jr, J. (1996). Entropically driven binding in a self-assembling molecular capsule. *Nature*, *382*(6588), 239.

Kästner, J. (2011). Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *1*(6), 932–942.

Kavitha, T., & Velraj, G. (2017). Density functional theory analysis and molecular docking evaluation of 1-(2, 5-dichloro-4-sulfophenyl)-3-methyl-5-pyrazolone as cox2 inhibitor against inflammatory diseases. *Journal of Molecular Structure*, *1141*, 335–345.

Kawaguchi, K., Saito, H., & Nagao, H. (2016). Decomposition analysis of free energy profile for hsp90-adp association. *Molecular Simulation*, *42*(11), 896–901.

Kawaguchi, K., Saito, H., Okazaki, S., & Nagao, H. (2013). Molecular dynamics study on the free energy profile for dissociation of adp from n-terminal domain of hsp90. *Chemical Physics Letters*, *588*, 226–230.

Kim, H. H., Hyun, J. S., Choi, J., Choi, K. E., Jee, J. G., & Park, S. J. (2018). Structural ensemble-based docking simulation and biophysical studies discovered new inhibitors of hsp90 n-terminal domain. *Scientific Reports*, *8*(1), 368.

Kirkwood, J. G. (1935). Statistical mechanics of fluid mixtures. *The Journal of Chemical Physics*, *3*(5), 300–313.

Kokh, D. B., Wade, R. C., & Wenzel, W. (2011). Receptor flexibility in small-molecule docking calculations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *1*(2), 298–314.

Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., & Wang, W. (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of Chemical Research*, *33*(12), 889–897.

Kramer, B., Rarey, M., & Lengauer, T. (1999). Evaluation of the flexx incremental construction algorithm for protein–ligand docking. *Proteins: Structure, Function, and Bioinformatics*, *37*(2), 228–241.

Krukenberg, K. A., Street, T. O., Lavery, L. A., & Agard, D. A. (2011). Conformational dynamics of the molecular chaperone hsp90. *Quarterly Reviews of Biophysics*, *44*(2), 229–255.

Kuhn, B., & Kollman, P. A. (2000). Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. *Journal of Medicinal Chemistry*, *43*(20), 3786–3791.

Kumar, A., & Zhang, K. Y. (2013). Investigation on the effect of key water molecules on docking performance in csardock exercise. *Journal of Chemical Information and Modeling*, *53*(8), 1880–1892.

Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, *13*(8), 1011–1021.

Kumari, R., Kumar, R., Consortium, O. S. D. D., & Lynn, A. (2014). g_mmpbsa a gromacs tool for high-throughput mm-pbsa calculations. *Journal of Chemical Information and Modeling*, *54*(7), 1951–1962.

Laio, A., & Parrinello, M. (2002). Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, *99*(20), 12562–12566.

Levitt, M., & Warshel, A. (1975). Computer simulation of protein folding. *Nature*, *253*(5494), 694.

Levy, Y., Jortner, J., & Becker, O. M. (2001). Solvent effects on the energy landscapes and folding kinetics of polyalanine. *Proceedings of the National Academy of Sciences*, *98*(5), 2188–2193.

Li, Y., Zhang, T., Schwartz, S. J., & Sun, D. (2009). New developments in hsp90 inhibitors as anti-cancer therapeutics: mechanisms, clinical perspective and more potential. *Drug Resistance Updates*, *12*(1-2), 17–27.

Limongelli, V., Bonomi, M., & Parrinello, M. (2013). Funnel metadynamics as accurate binding free-energy method. *Proceedings of the National Academy of Sciences*, *110*(16), 6358–6363.

Lin, J. H., Perryman, A. L., Schames, J. R., & McCammon, J. A. (2002). Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *Journal of the American Chemical Society*, *124*(20), 5632–5633.

Lin, J. H., Perryman, A. L., Schames, J. R., & McCammon, J. A. (2003). The relaxed complex method: Accommodating receptor flexibility for drug design with an improved scoring scheme. *Biopolymers: Original Research on Biomolecules*, *68*(1), 47–62.

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., & Shaw, D. E. (2010). Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins: Structure, Function, and Bioinformatics*, *78*(8), 1950–1958.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, *23*(1-3), 3–25.

MacKerell Jr, A. D., Banavali, N., & Foloppe, N. (2000). Development and current status of the charmm force field for nucleic acids. *Biopolymers: Original Research on Biomolecules*, *56*(4), 257–265.

Mbatha, S. H. (2015). *Identification of anti-cancer agents using integrated computational tools.* PhD thesis.

McCammon, J. A., Gelin, B. R., & Karplus, M. (1977). Dynamics of folded proteins. *Nature*, *267*(5612), 585.

Meyer, P., Prodromou, C., Hu, B., Vaughan, C., Roe, S. M., Panaretou, B., Piper, P. W., & Pearl, L. H. (2003). Structural and functional analysis of the middle segment of hsp90: implications for atp hydrolysis and client protein and cochaperone interactions. *Molecular Cell*, *11*(3), 647–658.

Michel, J., & Cuchillo, R. (2012). The impact of small molecule binding on the energy landscape of the intrinsically disordered protein c-myc. *PloS One*, *7*(7), e41070.

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., & Olson, A. J. (1998). Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, *19*(14), 1639–1662.

Nosé, S. (1986). An extension of the canonical ensemble molecular dynamics method. *Molecular Physics*, *57*(1), 187–191.

Ntie-Kang, F., Mbah, J. A., Lifongo, L. L., Owono, L. C. O., Megnassan, E., Mbaze, L. M., Judson, P. N., Sippl, W., & Efange, S. M. (2013). Assessing the pharmacokinetic profile of the cammednp natural products database: an in silico approach. *Organic and Medicinal Chemistry Letters*, *3*(1), 10.

Nyandoro, S. S., Ndanu, J., Munissi, J. J., Gruhonjic, A., Fitzpatrick, P. A., Landberg, G., Lu, Y., Wang, B., Pan, F., & Rissanen, K. (2015). N-cinnamoyltetraketide derivatives from the leaves of toussaintia orientalis. *Journal of Natural Products*, *78*(8), 2045–2050.

Oda, A., Okayasu, M., Kamiyama, Y., Yoshida, T., Takahashi, O., & Matsuzaki, H. (2007). Evaluation of docking accuracy and investigations of roles of parameters and each term in scoring functions for protein–ligand docking using arguslab software. *Bulletin of the Chemical Society of Japan*, *80*(10), 1920–1925.

Oostenbrink, C., Villa, A., Mark, A. E., & Van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the gromos force-field parameter sets 53a5 and 53a6. *Journal of Computational Chemistry*, *25*(13), 1656–1676.

Papadakis, R., & Deligkiozi, I. (2019). Solvent effects in supramolecular systems. In *Solvents and Solvent Effects*. IntechOpen.

Parrinello, M., & Rahman, A. (1980). Crystal structure and pair potentials: A molecular-dynamics study. *Physical Review Letters*, *45*(14), 1196.

Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, *52*(12), 7182–7190.

Patel, P. D., Yan, P., Seidler, P. M., Patel, H. J., Sun, W., Yang, C., Que, N. S., Taldone, T., Finotti, P., & Stephani, R. A. (2013). Paralog-selective hsp90 inhibitors define tumor-specific regulation of her2. *Nature Chemical Biology*, *9*(11), 677–684.

Patsahan, T., Ilnytskyi, J., & Pizio, O. (2017). On the properties of a single opls-ua model curcumin molecule in water, methanol and dimethyl sulfoxide. molecular dynamics computer simulation results. *arXiv preprint arXiv:1706.07253*.

Peräkylä, M., & Nordman, N. (2001). Energetic analysis of binding of progesterone and 5β-androstane-3, 17-dione to anti-progesterone antibody db3 using molecular dynamics and free energy calculations. *Protein Engineering*, *14*(10), 753–758.

Peterson, L. B. (2012). *Investigation of the Hsp90 C-terminal binding site, novel inhibitors and isoform-dependent client proteins*. PhD thesis, University of Kansas.

Philippova, O. E., Volkov, E. V., Sitnikova, N. L., Khokhlov, A. R., Desbrieres, J., & Rinaudo, M. (2001). Two types of hydrophobic aggregates in aqueous solutions of chitosan and its hydrophobic derivative. *Biomacromolecules*, *2*(2), 483–490.

Piaz, F., Vassallo, A., Chini, M. G., Cordero, F. M., Cardona, F., Pisano, C., Bifulco, G., De Tommasi, N., & Brandi, A. (2012). Natural iminosugar (+)-lentiginosine inhibits atpase and chaperone activity of hsp90. *PLoS One*, *7*(8), e43316.

Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., & Pearl, L. H. (1997). Identification and structural characterization of the atp/adp-binding site in the hsp90 molecular chaperone. *Cell*, *90*(1), 65–75.

Proia, D. A., Smith, D. L., Zhang, J., Jimenez, J. P., Sang, J., Ogawa, L. S., Sequeira, M., Acquaviva, J., He, S., & Zhang, C. (2015). Hsp90 inhibitor–sn-38 conjugate strategy for targeted delivery of topoisomerase i inhibitor to tumors. *Molecular Cancer Therapeutics*, *14*(11), 2422–2432.

Puccetti, L., Fasolis, G., Vullo, D., Chohan, Z. H., Scozzafava, A., & Supuran, C. T. (2005). Carbonic anhydrase inhibitors. inhibition of cytosolic/tumor-associated carbonic anhydrase isozymes i, ii, ix, and xii with schiffs bases incorporating chromone and aromatic sulfonamide moieties, and their zinc complexes. *Bioorganic & Medicinal Chemistry Letters*, *15*(12), 3096–3101.

Qin, H. L., Shang, Z. P., Jantan, I., Tan, O. U., Hussain, M. A., Sher, M., & Bukhari, S. N. A. (2015). Molecular docking studies and biological evaluation of chalcone based pyrazolines as tyrosinase inhibitors and potential anticancer agents. *RSC Advances*, *5*(57), 46330–46338.

Ramírez, D., & Caballero, J. (2016). Is it reliable to use common molecular docking methods for comparing the binding affinities of enantiomer pairs for their protein target? *International Journal of Molecular Sciences*, *17*(4), 525.

Rampino, A., Borgogna, M., Blasi, P., Bellich, B., & Cesàro, A. (2013). Chitosan nanoparticles: preparation, size evolution and stability. *International journal of pharmaceutics*, *455*(1-2), 219–228.

Rao, S. N., Singh, U. C., Bash, P. A., & Kollman, P. A. (1987). Free energy perturbation calculations on binding and catalysis after mutating asn 155 in subtilisin. *Nature*, *328*(6130), 551.

Rarey, M., Kramer, B., & Lengauer, T. (1999). The particle concept: placing discrete water molecules during protein-ligand docking predictions. *Proteins: Structure, Function, and Bioinformatics*, *34*(1), 17–28.

Rastelli, G., Rio, A. D., Degliesposti, G., & Sgobba, M. (2010). Fast and accurate predictions of binding free energies using mm-pbsa and mm-gbsa. *Journal of Computational Chemistry*, *31*(4), 797–810.

Razmimanesh, F., Amjad-Iranagh, S., & Modarress, H. (2015). Molecular dynamics simulation study of chitosan and gemcitabine as a drug delivery system. *Journal of Molecular Modeling*, *21*(7), 165.

Ribeiro, J. M. L., Tsai, S. T., Pramanik, D., Wang, Y., & Tiwary, P. (2018). Ligand dissociation mechanisms from all-atom simulations: Are we there yet? *arXiv preprint arXiv:1809.04540*.

Roberts, B. C., & Mancera, R. L. (2008). Ligand-protein docking with water molecules. *Journal of Chemical Information and Modeling*, *48*(2), 397–408.

Roe, S. M., Prodromou, C., O'Brien, R., Ladbury, J. E., Piper, P. W., & Pearl, L. H. (1999). Structural basis for inhibition of the hsp90 molecular chaperone by the antitumor antibiotics radicicol and geldanamycin. *Journal of Medicinal Chemistry*, *42*(2), 260–266.

Rungnim, C., Rungrotmongkol, T., Hannongbua, S., & Okumura, H. (2013). Replica exchange molecular dynamics simulation of chitosan for drug delivery system based on carbon nanotube. *Journal of Molecular Graphics and Modelling*, *39*, 183–192.

Ruppert, J., Welch, W., & Jain, A. N. (1997). Automatic identification and representation of protein binding sites for molecular docking. *Protein Science*, *6*(3), 524–533.

Rwegasila, E., Munissi, J. J., Mubofu, E., Nyandoro, S. S., & Erasto, P. (2018). In vivo antimycobacterial studies of toussaintine a-chitosan nanocomposites. *Tanzania Journal of Science*, *44*(2), 16–26.

Santos, R., Hritz, J., & Oostenbrink, C. (2009). Role of water in molecular docking simulations of cytochrome p450 2d6. *Journal of Chemical Information and Modeling*, *50*(1), 146–154.

Sanyakamdhorn, S., Agudelo, D., & Tajmir-Riahi, H.-A. (2013). Encapsulation of antitumor drug doxorubicin and its analogue by chitosan nanoparticles. *Biomacromolecules*, *14*(2), 557–563.

Sauton, N., Lagorce, D., Villoutreix, B. O., & Miteva, M. A. (2008). Ms-dock: accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening. *BMC Bioinformatics*, *9*(1), 184.

Sauvage, F., Messaoudi, S., Fattal, E., Barratt, G., & Vergnaud-Gauduchon, J. (2017). Heat shock proteins and cancer: How can nanomedicine be harnessed? *Journal of Controlled Release*, *248*, 133–143.

Saxena, A., Saxena, S., & Chaudhaery, S. (2010). Molecular modelling and docking studies on heat shock protein 90 (hsp90) inhibitors. *SAR and QSAR in Environmental Research*, *21*(1-2), 1–20.

Schames, J. R., Henchman, R. H., Siegel, J. S., Sotriffer, C. A., Ni, H., & McCammon, J. A. (2004). Discovery of a novel binding trench in hiv integrase. *Journal of Medicinal Chemistry*, *47*(8), 1879–1881.

Schmid, N., Eichenberger, A. P., Choutko, A., Riniker, S., Winger, M., Mark, A. E., & van Gunsteren, W. F. (2011). Definition and testing of the gromos force-field versions 54a7 and 54b7. *European Biophysics Journal*, *40*(7), 843.

Schneider, N., Lange, G., Hindle, S., Klein, R., & Rarey, M. (2013). A consistent description of hydrogen bond and dehydration energies in protein–ligand complexes: methods behind the hyde scoring function. *Journal of Computer-Aided Molecular Design*, *27*(1), 15–29.

Schüttelkopf, A. W., & Van Aalten, D. M. (2004). Prodrg: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*, *60*(8), 1355–1363.

Sepehri, B., & Ghavami, R. (2018). Towards the in-silico design of new hsp90 inhibitors: Molecular docking and 3d-qsar comfa studies of tetrahydropyrido [4, 3-d] pyrimidine derivatives as hsp90 inhibitors. *Medicinal Chemistry*, *14*(5), 439–450.

Sha, K., & Cao, W. (2015). Structural and energetic insight into the isoform-selective inhibitors of tumour marker hsp90 against grp94. *Molecular Simulation*, *41*(18), 1553–1561.

Shadrack, D. M., Swai, H. S., Munissi, J. J., Mubofu, E. B., & Nyandoro, S. S. (2018). Polyamidoamine dendrimers for enhanced solubility of small molecules and other desirable properties for site specific delivery: Insights from experimental and computational studies. *Molecules*, *23*(6), 1419.

Shamsara, J. (2018). Correlation between virtual screening performance and binding site descriptors of protein targets. *International Journal of Medicinal Chemistry*, *2018*.

Shan, P., Shen, J. W., Xu, D. H., Shi, L. Y., Gao, J., Lan, Y. W., Wang, Q., & Wei, X. H. (2014). Molecular dynamics study on the interaction between doxorubicin and hydrophobically modified chitosan oligosaccharide. *RSC Advances*, *4*(45), 23730–23739.

Sheehan, M. E., & Sharratt, P. N. (1998). Molecular dynamics methodology for the study of the solvent effects on a concentrated diels-alder reaction and the separation of the post-reaction mixture. *Computers & Chemical Engineering*, *22*, S27–S33.

Shen, J. W., Li, J., Zhao, Z., Zhang, L., Peng, G., & Liang, L. (2017). Molecular dynamics study on the mechanism of polynucleotide encapsulation by chitosan. *Scientific Reports*, *7*(1), 5050.

Shi, J., Van de Water, R., Hong, K., Lamer, R. B., Weichert, K. W., Sandoval, C. M., Kasibhatla, S. R., Boehm, M. F., Chao, J., & Lundgren, K. (2012). Ec144 is a potent inhibitor of the heat shock protein 90. *Journal of Medicinal Chemistry*, *55*(17), 7786–7795.

Shirts, M. R., & Pande, V. S. (2005). Comparison of efficiency and bias of free energies computed by exponential averaging, the bennett acceptance ratio, and thermodynamic integration. *The Journal of Chemical Physics*, *122*(14), 144107.

Simunovic, M., & Voth, G. A. (2012). Molecular and thermodynamic insights into the conformational transitions of hsp90. *Biophysical Journal*, *103*(2), 284–292.

Sinko, W., Lindert, S., & McCammon, J. A. (2013). Accounting for receptor flexibility and enhanced sampling methods in computer-aided drug design. *Chemical Biology & Drug Design*, *81*(1), 41–49.

Slabber, C. A., Grimmer, C. D., & Robinson, R. S. (2016). Solution conformations of curcumin in dmso. *Journal of Natural Products*, *79*(10), 2726–2730.

Slakman, B. L., & West, R. H. (2019). Kinetic solvent effects in organic reactions. *Journal of Physical Organic Chemistry*, *32*(3), 3904.

Smithrud, D. B., & Diederich, F. (1990). Strength of molecular complexation of apolar solutes in water and in organic solvents is predictable by linear free energy relationships: A general model for solvation effects on apolar binding. *Journal of the American Chemical Society*, *112*(1), 339–343.

Soga, S., Akinaga, S., & Shiotsu, Y. (2013). Hsp90 inhibitors as anti-cancer agents, from basic discoveries to clinical development. *Current Pharmaceutical Design*, *19*(3), 366–376.

Sohraby, F., Bagheri, M., & Aryapour, H. (2019). Performing an in silico repurposing of existing drugs by combining virtual screening and molecular dynamics simulation. In *Computational Methods for Drug Repurposing*. Springer.

Soldano, K. L., Jivan, A., Nicchitta, C. V., & Gewirth, D. T. (2003). Structure of the n-terminal domain of grp94 basis for ligand specificity and regulation. *Journal of Biological Chemistry*, *278*(48), 48330–48338.

Stebbins, C. E., Russo, A. A., Schneider, C., Rosen, N., Hartl, F. U., & Pavletich, N. P. (1997). Crystal structure of an hsp90–geldanamycin complex: targeting of a protein chaperone by an antitumor agent. *Cell*, *89*(2), 239–250.

Stjernschantz, E., & Oostenbrink, C. (2010). Improved ligand-protein binding affinity predictions using multiple binding modes. *Biophysical Journal*, *98*(11), 2682–2691.

Su, P. C., & Johnson, M. E. (2016). Evaluating thermodynamic integration performance of the new amber molecular dynamics package and assess potential halogen bonds of enoyl-acp reductase (fabi) benzimidazole inhibitors. *Journal of Computational Chemistry*, *37*(9), 836–847.

Sun, H., Li, Y., Tian, S., Xu, L., & Hou, T. (2014). Assessing the performance of mm/pbsa and mm/gbsa methods. 4. accuracies of mm/pbsa and mm/gbsa methodologies evaluated by various simulation protocols using pdbbind data set. *Physical Chemistry Chemical Physics*, *16*(31), 16719–16729.

Sutherland, J. J., Nandigam, R. K., Erickson, J. A., & Vieth, M. (2007). Lessons in molecular recognition. 2. assessing and improving cross-docking accuracy. *Journal of Chemical Information and Modeling*, *47*(6), 2293–2302.

Swope, W. C., Andersen, H. C., Berens, P. H., & Wilson, K. R. (1982). A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, *76*(1), 637–649.

Swuec, P., & Barlow, D. J. (2012). Prediction of inhibitory activities of hsp90 inhibitors. *Bioorganic & Medicinal Chemistry*, *20*(1), 408–414.

Thilagavathi, R., & Mancera, R. L. (2010). Ligand- protein cross-docking with water molecules. *Journal of Chemical Information and Modeling*, *50*(3), 415–421.

Tian, Z. Q., Liu, Y., Zhang, D., Wang, Z., Dong, S. D., Carreras, C. W., Zhou, Y., Rastelli, G., Santi, D. V., & Myles, D. C. (2004). Synthesis and biological activities of novel 17-aminogeldanamycin derivatives. *Bioorganic & Medicinal Chemistry*, *12*(20), 5317–5329.

Tiwary, P., & Parrinello, M. (2014). A time-independent free energy estimator for metadynamics. *The Journal of Physical Chemistry B*, *119*(3), 736–742.

Totrov, M., & Abagyan, R. (2008). Flexible ligand docking to multiple receptor conformations: a practical alternative. *Current Opinion in Structural Biology*, *18*(2), 178–184.

Tovchigrechko, A., & Vakser, I. A. (2006). Gramm-x public web server for protein–protein docking. *Nucleic Acids Research*, *34*(2), 310–314.

Trott, O., & Olson, A. J. (2010). Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461.

Van Gunsteren, W. F., & Berendsen, H. J. (1988). A leap-frog algorithm for stochastic dynamics. *Molecular Simulation*, *1*(3), 173–185.

Verdonk, M. L., Berdini, V., Hartshorn, M. J., Mooij, W. T., Murray, C. W., Taylor, R. D., & Watson, P. (2004). Virtual screening using protein- ligand docking: avoiding artificial enrichment. *Journal of Chemical Information and Computer Sciences*, *44*(3), 793–806.

Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., & Taylor, R. D. (2003). Improved protein–ligand docking using gold. *Proteins: Structure, Function, and Bioinformatics*, *52*(4), 609–623.

Verlet, L. (1967). Computer" experiments" on classical fluids. I. thermodynamical properties of lennard-jones molecules. *Physical Review*, *159*(1), 98.

Vijesh, A., Isloor, A. M., Telkar, S., Arulmoli, T., & Fun, H. K. (2013). Molecular docking studies of some new imidazole derivatives for antimicrobial properties. *Arabian Journal of Chemistry*, *6*(2), 197–204.

Wang, Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, *25*(9), 1157–1174.

Wang, M., Shen, G., & Blagg, B. S. (2006). Radanamycin, a macrocyclic chimera of radicicol and geldanamycin. *Bioorganic & Medicinal Chemistry Letters*, *16*(9), 2459–2462.

Wang, X. Y., Zhang, L., Wei, X. H., & Wang, Q. (2013). Molecular dynamics of paclitaxel encapsulated by salicylic acid-grafted chitosan oligosaccharide aggregates. *Biomaterials*, *34*(7), 1843–1851.

Warren, G. L., Andrews, C. W., Capelli, A. M., Clarke, B., LaLonde, J., Lambert, M. H., Lindvall, M., Nevins, N., Semus, S. F., & Senger, S. (2006). A critical assessment of docking programs and scoring functions. *Journal of Medicinal Chemistry*, *49*(20), 5912–5931.

Weis, A., Katebzadeh, K., Söderhjelm, P., Nilsson, I., & Ryde, U. (2006). Ligand affinities predicted with the mm/pbsa method: dependence on the simulation method and the force field. *Journal of Medicinal Chemistry*, *49*(22), 6596–6606.

Wishart, D. S., Knox, C., Guo, A. C., Cheng, D., Shrivastava, S., Tzur, D., Gautam, B., & Hassanali, M. (2007). Drugbank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Research*, *36*(1), 901–906.

Xu, L., Sun, H., Li, Y., Wang, J., & Hou, T. (2013). Assessing the performance of mm/pbsa and mm/gbsa methods. 3. the impact of force fields and ligand charge models. *The Journal of Physical Chemistry B*, *117*(28), 8408–8421.

Yan, A., Grant, G. H., & Richards, W. G. (2008). Dynamics of conserved waters in human hsp90: implications for drug design. *Journal of The Royal Society Interface*, *5*(3), 199–205.

Ytreberg, F. M., Swendsen, R. H., & Zuckerman, D. M. (2006). Comparison of free energy methods for molecular systems. *The Journal of Chemical Physics*, *125*(18), 184114.

Yun, T. J., Harning, E. K., Giza, K., Rabah, D., Li, P., Arndt, J. W., Luchetti, D., Biamonte, M. A., Shi, J., & Lundgren, K. (2011). Ec144, a synthetic inhibitor of heat shock protein 90, blocks innate and adaptive immune responses in models of inflammation and autoimmunity. *The Journal of Immunology*, *186*(1), 563–575.

Curcumin in solvents.

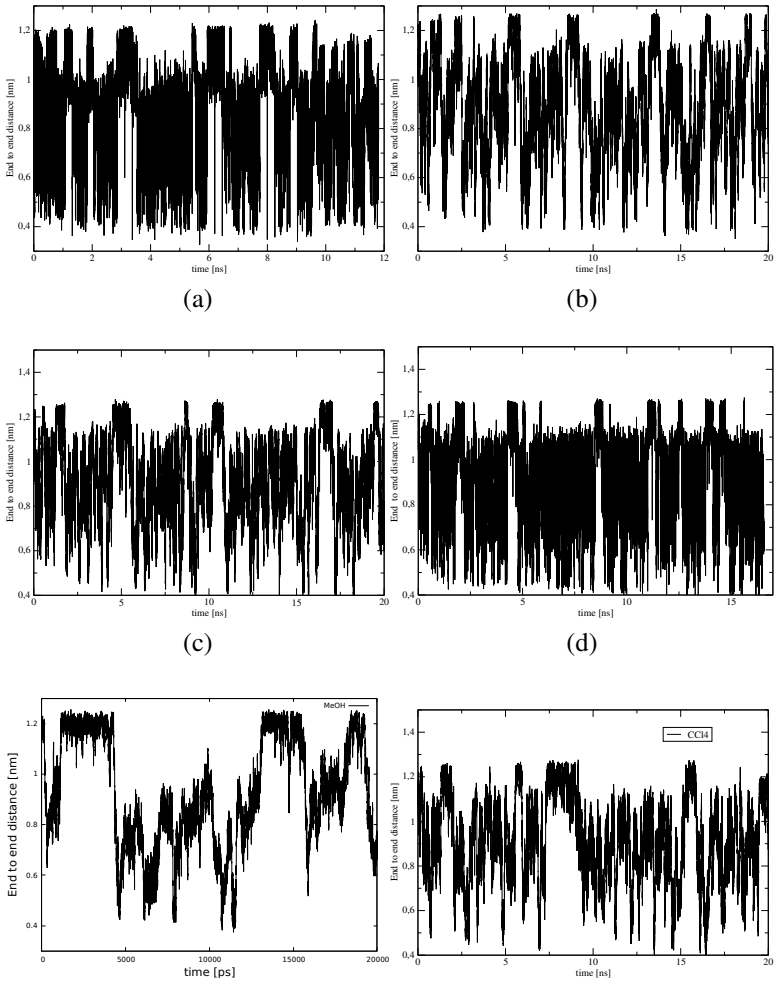
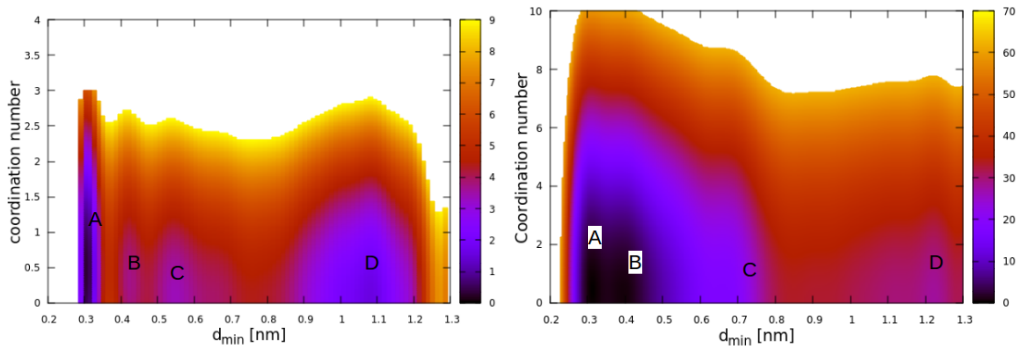


Figure 36: Time dependence for end to end distance in different solvents, (a) vacuum, (b) water, (c) DMSO, (d) DCM, (e) MeOH and (f) CCl4.

Chitosan-TouA supramolecular interaction.



(a) DMSO (b) Water

Figure 37: 2D free energy as a function of minimum distance and coordination number in two solvents. In both solvents, minimum A and B corresponds to bound conformation while minimum D corresponds the unbound conformations. In DMSO large free energy for unbound is observed compared to the system in water.

Thermodynamics integration.

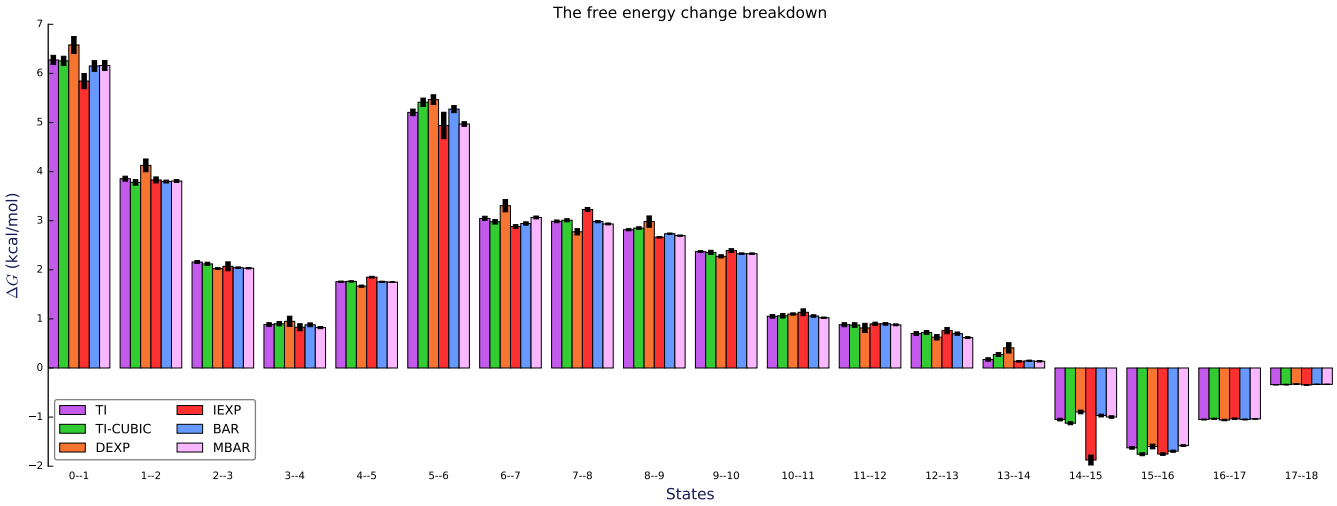


Figure 38: Free energy change breakdown for chitosan-TouA complex in water

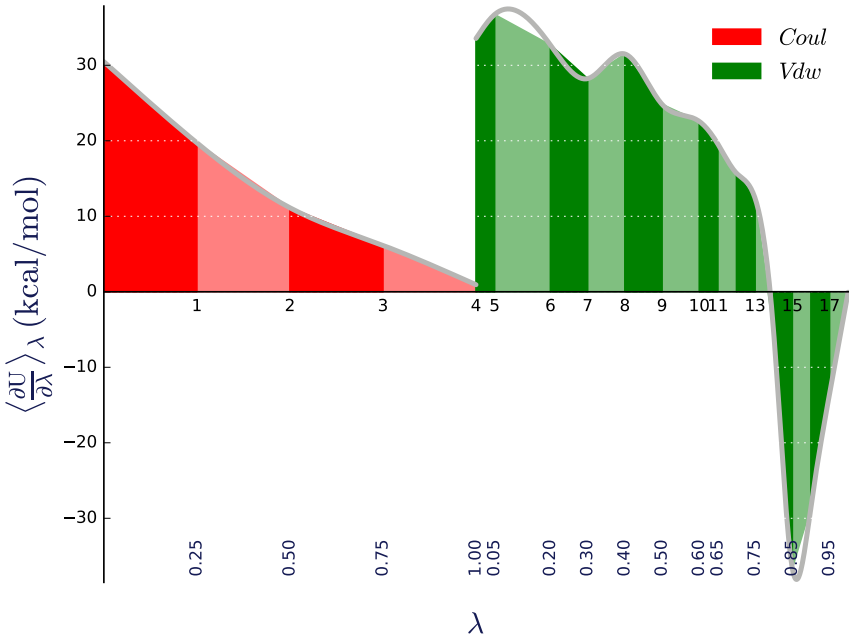


Figure 39: Free energy for chitosan-TouA complex in water

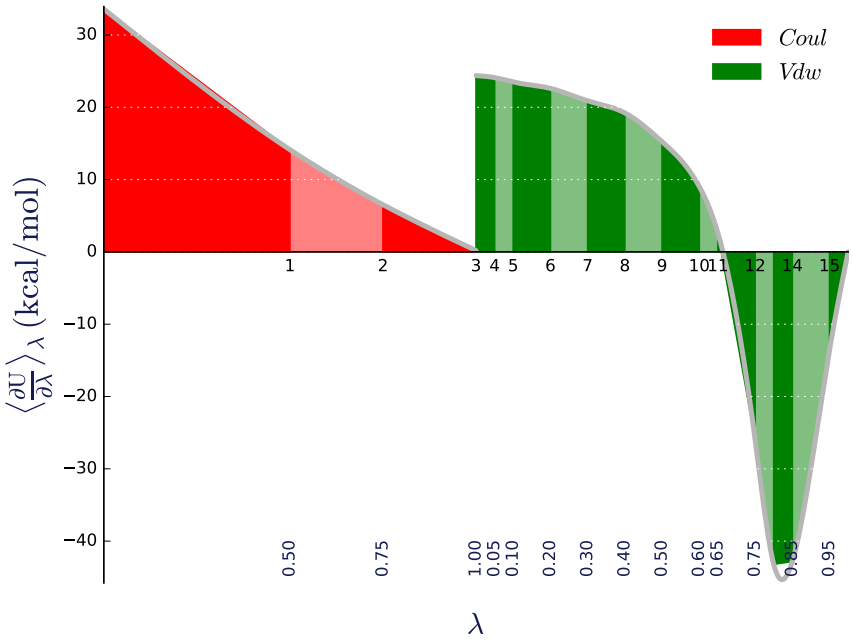


Figure 40: Free energy TouA complex in water

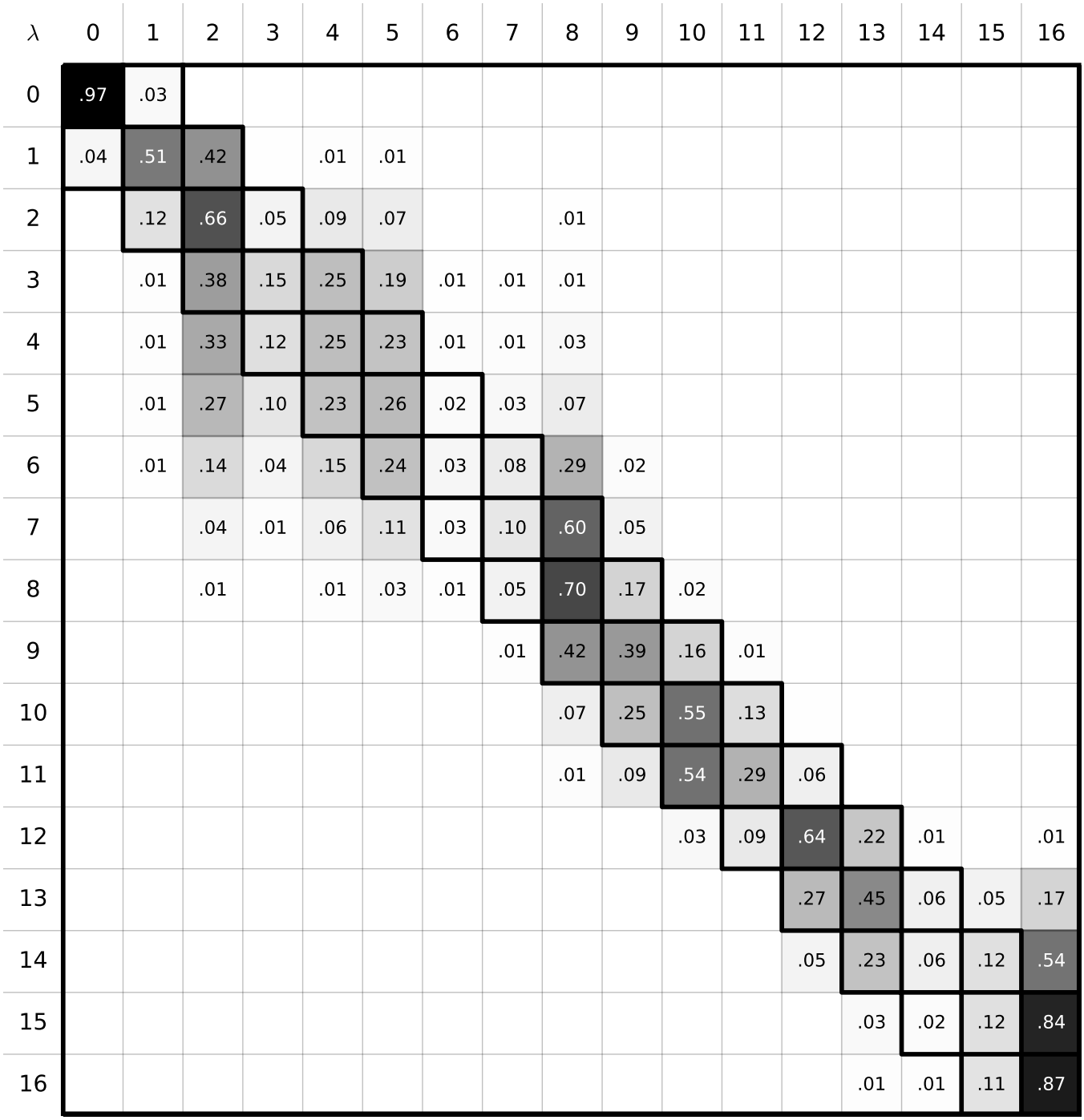


Figure 41: MBAR convergence of free energy for the complex in water